INSTITUTE OF PHYSICS PUBLISHING REPORTS ON PROGRESS IN PHYSICS

Rep. Prog. Phys. **68** (2005) 2267–2336 doi:10.1088/0034-4885/68/10/R01

HgCdTe infrared detector material: history, status and outlook

A Rogalski

Institute of Applied Physics, Military University of Technology, 2 Kaliskiego St., 00-908 Warsaw, Poland

Received 10 November 2004, in final form 13 April 2005
Published 22 August 2005
Online at stacks.iop.org/RoPP/68/2267

Abstract

This article reviews the history, the present status and possible future developments of HgCdTe ternary alloy for infrared (IR) detector applications. HgCdTe IR detectors have been intensively developed since the first synthesis of this material in 1958. This article summarizes the fundamental properties of this versatile narrow gap semiconductor, and relates the material properties to its successful applications as an IR photoconductive and photovoltaic detector material. An emphasis is put on key developments in the crystal growth and their influence on device evolution. Competitive technologies to HgCdTe ternary alloy are also presented.

Recent advances of backside illuminated HgCdTe heterojunction photodiodes have enabled a third generation of multispectral instruments for remote sensing applications and have led to the practicality of multiband IR focal plane array technology. Finally, evaluation of HgCdTe for room temperature long wavelength IR applications is presented.

(Some figures in this article are in colour only in the electronic version)

Contents

	Page
1. Introduction	2269
2. Fundamental HgCdTe properties	2269
2.1. Energy band gap	2271
2.2. Mobilities	2272
2.3. Optical properties	2272
2.4. Thermal generation–recombination processes	2274
2.4.1. Shockley–Read processes	2274
2.4.2. Radiative processes	2275
2.4.3. Auger processes	2275
2.5. α/G ratio and detectivity	2276
3. Photodiodes—principle of operation and figure of merit	2279
4. Historical perspective	2284
5. Impact of epitaxial growth on development of HgCdTe detectors	2291
6. HgCdTe photodiodes	2296
7. Third generation detectors	2306
7.1. Noise equivalent difference temperature	2306
7.2. Pixel and chip sizes	2307
7.3. Uniformity	2308
7.4. Identification and detection ranges	2309
7.5. Two-colour HgCdTe detectors	2311
8. Alternative material systems	2315
8.1. Lead salt ternary alloys	2315
8.2. InSb and InGaAs	2316
8.3. GaAs/AlGaAs QWIPs	2319
8.4. InAs/GaInSb strained layer superlattices	2320
8.5. Hg-based alternatives to HgCdTe	2322
9. HgCdTe versus thermal detectors	2322
10. Summary	2330
References	2331

1. Introduction

In 1959, a publication by Lawson *et al* [1] triggered the development of variable band gap $Hg_{1-x}Cd_xTe$ (HgCdTe) alloys providing an unprecedented degree of freedom in infrared (IR) detector design. HgCdTe is a pseudo-binary alloy semiconductor that crystallizes in a zinc blende structure. Because of its band gap tunability with x, $Hg_{1-x}Cd_xTe$ has evolved to become the most important/versatile material for detector applications over the entire IR range. As the Cd composition increases, the energy gap for $Hg_{1-x}Cd_xTe$ gradually increases from a negative value for HgTe to a positive value for CdTe. The band gap energy tunability results in IR detector applications that span the short wavelength IR (SWIR: 1–3 μm), middle wavelength (MWIR: 3–5 μm), long wavelength (LWIR: 8–14 μm) and very long wavelength (VLWIR: 14–30 μm) ranges.

HgCdTe technology development was and continues to be primarily for military applications. A negative aspect of support of defence agencies has been the associated secrecy requirements that inhibit meaningful collaborations among research teams on a national and especially on an international level. In addition, the primary focus has been on focal plane array (FPA) demonstration and much less on establishing the knowledge base. Nevertheless, significant progress has been made over four decades. At present HgCdTe is the most widely used variable gap semiconductor for IR photodetectors.

The specific advantages of HgCdTe are the direct energy gap, ability to obtain both low and high carrier concentrations, high mobility of electrons and low dielectric constant. The extremely small change of lattice constant with composition makes it possible to grow high quality layers and heterostructures. HgCdTe can be used for detectors operated at various modes, and can be optimized for operation at the extremely wide range of the IR spectrum (1–30 μm) and at temperatures ranging from that of liquid helium to room temperature.

The main motivations for replacing HgCdTe are the technological disadvantages of this material. One of them is a weak Hg–Te bond, which results in bulk, surface and interface instabilities. Uniformity and yield are still issues especially in the LWIR spectral range. Nevertheless, HgCdTe remains the leading semiconductor for IR detectors.

This article summarizes the fundamental properties of HgCdTe versatile alloy semiconductor, and relates these material properties, which have successful applications as an IR detector material. The intent is to concentrate on device approaches and technologies that are having the most impact today mainly in photovoltaic detectors. A secondary aim is to outline the historical evolution of HgCdTe detector technology showing why certain device designs and architecture have emerged as more successful and more useful today, especially in FPA technology and development of third generation IR detectors. Also alternative technologies competitive to HgCdTe, technology are considered.

2. Fundamental HgCdTe properties

HgCdTe ternary alloy is a nearly ideal IR detector material system. Its position is conditioned by three key features [2]:

- tailorable energy band gap over the 1–30 μm range,
- large optical coefficients that enable high quantum efficiency and
- favourable inherent recombination mechanisms that lead to high operating temperature.

These properties are a direct consequence of the energy band structure of the zinc blende semiconductor. Moreover, the specific advantages of HgCdTe are the ability to obtain both low and high carrier concentrations, high mobility of electrons and low dielectric

Table 1. Summary of the material properties for the $Hg_{1-x}Cd_xTe$ ternary alloy, listed for the binary components HgTe and CdTe, and for several technologically important alloy compositions (after [2]).

Property	HgTe	$Hg_{1-x}Cd_xTe$						CdTe
x	0	0.194	0.205	0.225	0.31	0.44	0.62	1.0
a (Å)	6.461	6.464	6.464	6.464	6.465	6.468	6.472	6.481
	77 K	77 K	77 K	77 K	140 K	200 K	250 K	300 K
E_g (eV)	−0.261	0.073	0.091	0.123	0.272	0.474	0.749	1.490
λ_c (μm)	—	16.9	13.6	10.1	4.6	2.6	1.7	0.8
n_i (cm^{-3})	—	1.9×10^{14}	5.8×10^{13}	6.3×10^{12}	3.7×10^{12}	7.1×10^{11}	3.1×10^{10}	4.1×10^{5}
m_c/m_0	—	0.006	0.007	0.010	0.021	0.035	0.053	0.102
g_c	—	−150	−118	−84	−33	−15	−7	−1.2
$\varepsilon_s/\varepsilon_0$	20.0	18.2	18.1	17.9	17.1	15.9	14.2	10.6
$\varepsilon_\infty/\varepsilon_0$	14.4	12.8	12.7	12.5	11.9	10.8	9.3	6.2
n_r	3.79	3.58	3.57	3.54	3.44	3.29	3.06	2.50
μ_e (cm^2 V^{-1} s^{-1})	—	4.5×10^5	3.0×10^5	1.0×10^5	—	—	—	—
μ_{hh} (cm^2 V^{-1} s^{-1})	—	450	450	450	—	—	—	—
$b = \mu_e/\mu_\eta$	—	1000	667	222	—	—	—	—
τ_R (μs)	—	16.5	13.9	10.4	11.3	11.2	10.6	2
τ_{A1} (μs)	—	0.45	0.85	1.8	39.6	453	4.75×10^3	
$\tau_{typical}$ (μs)	—	0.4	0.8	1	7	—	—	—
E_p (eV)				19				
Δ (eV)				0.93				
m_{hh}/m_0				0.40–0.53				
ΔE_v (eV)				0.35–0.55				

τ_R and τ_{A1} calculated for n-type HgCdTe with $N_d = 1 \times 10^{15}$ cm^{-3}. The last four material properties are independent of or relatively insensitive to alloy composition.

Table 2. Some physical properties of narrow gap semiconductors.

	E_g (eV)		n_i (cm^{-3})			μ_e (10^4 cm^2 V^{-1} s^{-1})		μ_h (10^4 cm^2 V^{-1} s^{-1})	
Material	77 K	300 K	77 K	300 K	ε	77 K	300 K	77 K	300 K
InAs	0.414	0.359	6.5×10^3	9.3×10^{14}	14.5	8	3	0.07	0.02
InSb	0.228	0.18	2.6×10^9	1.9×10^{16}	17.9	100	8	1	0.08
$In_{0.53}Ga_{0.47}As$	0.66	0.75	—	5.4×10^{11}	14.6	7	1.38	—	0.05
PbS	0.31	0.42	3×10^7	1.0×10^{15}	172	1.5	0.05	1.5	0.06
PbSe	0.17	0.28	6×10^{11}	2.0×10^{16}	227	3	0.10	3	0.10
PbTe	0.22	0.31	1.5×10^{10}	1.5×10^{16}	428	3	0.17	2	0.08
$Pb_{1-x}Sn_xTe$	0.1	0.1	3.0×10^{13}	2.0×10^{16}	400	3	0.12	2	0.08
$Hg_{1-x}Cd_xTe$	0.1	0.1	3.2×10^{13}	2.3×10^{16}	18.0	20	1	0.044	0.01
$Hg_{1-x}Cd_xTe$	0.25	0.25	7.2×10^8	2.3×10^{15}	16.7	8	0.6	0.044	0.01

constant. The extremely small change of lattice constant with composition makes it possible to grow high quality layered and graded gap structures. As a result, HgCdTe can be used for detectors operated at various modes (photoconductor, photodiode or metal–insulator–semiconductor (MIS) detector).

Table 1 summarizes various material properties of $Hg_{1-x}Cd_xTe$, while table 2 compares important parameters of HgCdTe with other narrow gap semiconductors used in IR detector fabrication.

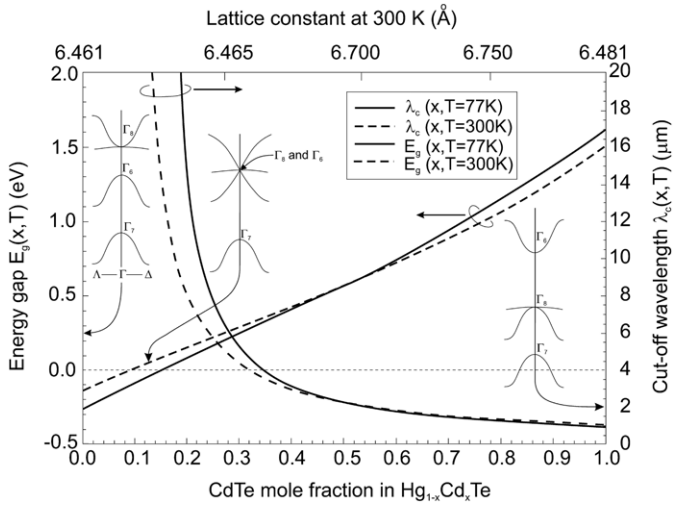


Figure 1. The band gap structure of $Hg_{1-x}Cd_xTe$ near the Γ-point for three different values of the forbidden energy gap. The energy band gap is defined at the difference between the Γ_6 and Γ_8 band extrema at $\Gamma = 0$.

2.1. Energy band gap

The electrical and optical properties of $Hg_{1-x}Cd_xTe$ are determined by the energy gap structure in the vicinity of the Γ-point of the Brillouin zone, in essentially the same way as InSb. The shape of the electron band and the light-mass hole band are determined by the $\boldsymbol{k}\cdot\boldsymbol{p}$ interaction, and hence, by the energy gap and the momentum matrix element. The energy gap of this compound at 4.2 K ranges from −0.300 eV for semimetallic HgTe, goes through zero at about $x = 0.15$ and opens up to 1.648 eV for CdTe.

Figure 1 plots the energy band gap $E_g(x, T)$ for $Hg_{1-x}Cd_xTe$ versus alloy composition parameter x at temperatures 77 and 300 K. Also plotted is the cutoff wavelength $\lambda_c(x, T)$, defined as that wavelength at which the response has dropped to 50% of its peak value.

A number of expressions approximating $E_g(x, T)$ are available at present [3]. The most widely used expression is due to Hansen *et al* [4]

$$E_g = -0.302 + 1.93x - 0.81x^2 + 0.832x^3 + 5.35 \times 10^{-4}(1 - 2x)T, \qquad (1)$$

where E_g is in electronvolt and T is in kelvin.

The expression which has become the most widely used for intrinsic carrier concentration is that of Hansen and Schmit [5] who used their own $E_g(x, T)$ relationship of (1), the $\boldsymbol{k}\cdot\boldsymbol{p}$ method and a value of $0.443\, m_0$ for heavy hole effective mass ratio

$$n_i = (5.585 - 3.82x + 0.001\,753T - 0.001\,364xT) \times 10^{14} E_g^{3/4} T^{3/2} \exp\left(-\frac{E_g}{2kT}\right). \qquad (2)$$

The electron m_e^* and light hole m_{lh}^* effective masses in the narrow gap mercury compounds are close and they can be established according to the Kane band model. Here we used Weiler's [6] expression

$$\frac{m_0}{m_e^*} = 1 + 2F + \frac{E_p}{3}\left(\frac{2}{E_g} + \frac{1}{E_g + \Delta}\right), \qquad (3)$$

where $E_p = 19$ eV, $\Delta = 1$ eV and $F = -0.8$. This relationship can be approximated by $m_e^*/m \approx 0.071\, E_g$, where E_g is in electronvolts. The effective mass of heavy hole m_{hh}^* is high;

the measured values range between 0.3 and $0.7m_0$. The value of $m^*_{hh} = 0.55\,m_0$ is frequently used in modelling of IR detectors.

2.2. Mobilities

Due to small effective masses the electron mobilities in HgCdTe are remarkably high, while heavy-hole mobilities are two orders of magnitude lower. A number of scattering mechanisms dominate the electron mobility [7, 8]. The x-dependence of the mobility results primarily from the x-dependence of the band gap, and the temperature dependence primarily from the competition among various scattering mechanisms which are temperature dependent.

The transport properties of holes are less studied than those of electrons mainly because the contribution of holes to the electrical conduction is relatively small due to their low mobility. Very often in transport measurements, the electron contribution predominates even in p-type materials unless the electron density is sufficiently low. Comprehensive analysis of different hole scattering mechanisms in $Hg_{1-x}Cd_xTe$ ($x = 0.2$–0.4) has been carried out by Yadava *et al* [9].

The electron mobility in $Hg_{1-x}Cd_xTe$ (expressed in $cm^2\,V^{-1}\,s^{-1}$), in composition range $0.2 \leqslant x \leqslant 0.6$ and temperature range $T > 50$ K, can be approximated as [10]:

$$\mu_e = \frac{9 \times 10^8 s}{T^{2r}}, \qquad \text{where } r = \left(\frac{0.2}{x}\right)^{0.6} \quad \text{and} \quad s = \left(\frac{0.2}{x}\right)^{7.5}. \tag{4}$$

Higgins *et al* [11] give an empirical formula (valid for $0.18 \leqslant x \leqslant 0.25$) for the variation of μ_e with x at 300 K for the very high quality melt grown samples that they studied:

$$\mu_e = 10^4(8.754x - 1.044)^{-1} \quad \text{in cm}^2(\text{V s})^{-1}. \tag{5}$$

The hole mobilities at room temperature range from 40 to $80\,cm^2(V\,s)^{-1}$, and the temperature dependence is relatively weak. A 77 K hole mobility is by one order of magnitude higher. According to [12], the hole mobility measured at 77 K falls as the acceptor concentration is increased and in the composition range 0.20–0.30 yields the following empirical expression:

$$\mu_h = \mu_0 \left[1 + \left(\frac{p}{1.8 \times 10^{17}}\right)^2\right]^{-1/4}, \tag{6}$$

where $\mu_0 = 440\,cm^2(V\,s)^{-1}$.

For modelling IR photodetectors, the hole mobility is usually calculated assuming that the electron-to-hole mobility ratio $b = \mu_e/\mu_h$ is constant and equal to 100.

2.3. Optical properties

Optical properties of HgCdTe have been mainly investigated at energies near the band gap [7, 13]. There still appears to be considerable disagreement among the reported results concerning absorption coefficients. This is caused by different concentrations of native defects and impurities, non-uniform composition and doping, thickness inhomogeneities of samples, mechanical strains and different surface treatments.

In high-quality samples the measured absorption in the short-wavelength region is in good agreement with the Kane model calculation, while the situation appears to be complicated in the long-wavelength edge by the appearance of an absorption tail extending at energies lower than

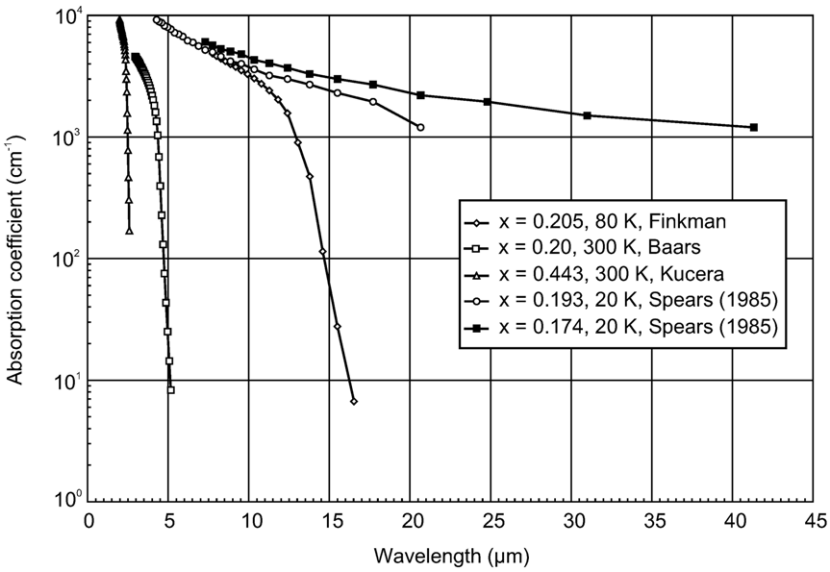


Figure 2. Optical absorption coefficient data for several $Hg_{1-x}Cd_xTe$ alloy compositions, for photon energies near the fundamental absorption edge, plotted versus wavelength (after [2]).

the energy gap. This tail has been attributed to the composition-induced disorder. According to Finkman and Schacham [14], the absorption tail obeys a modified Urbach's rule:

$$\alpha = \alpha_0 \exp\left[\frac{\sigma(E - E_0)}{T + T_0}\right] \quad \text{in cm}^{-1}, \tag{7}$$

where T is in kelvin, E is in electronvolts and $\alpha_0 = \exp(53.61x–18.88)$, $E_0 = -0.3424 + 1.838x + 0.148x^2$ (in electronvolts), $T_0 = 81.9$ (in kelvin), $\sigma = 3.267 \times 10^4(1 + x)$ (in kelvin per electronvolts) are fitting parameters which vary smoothly with composition. The fit was performed with data at $x = 0.215$ and $x = 1$ and for temperatures between 80 and 300 K.

Chu *et al* [15] have also found an empirical formula for the calculation of the intrinsic optical absorption coefficient at the Kane region

$$\alpha = \alpha_g \exp\left[\beta(E - E_g)\right]^{1/2}, \tag{8}$$

where the parameter β depends on the alloy composition and temperature: $\beta(x, T) = -1 + 0.083\,T + (21 - 0.13\,T)x$. Expanding equation (8) one finds a linear term, $(E - E_g)^{1/2}$, that fits the square-root law between α and E proper for parabolic bands.

Optical absorption coefficient data for several alloy compositions of $Hg_{1-x}Cd_xTe$ versus wavelength are plotted in figure 2. We can see that the absorption strength generally decreases as the gap becomes smaller, due both to the decrease in the conduction band effective mass and to the $\lambda^{-1/2}$ dependence of the absorption coefficient on wavelength λ.

The dielectric constants are not a linear function of x and temperature dependence was not observed within the experimental resolution [16]. These dependences can be described by the following relations:

$$\varepsilon_\infty = 15.2 - 15.6x + 8.2x^2, \tag{9}$$

$$\varepsilon_0 = 20.5 - 15.6x + 5.7x^2. \tag{10}$$

2.4. Thermal generation–recombination processes

The generation processes which compete against the recombination processes directly affect the performance of photodetectors, setting up a steady-state concentration of carriers in semiconductors subjected to thermal and optical excitation and, frequently determining the kinetics of photogenerated signals. Generation–recombination (g–r) processes in HgCdTe are widely discussed in the literature (see, e.g. [3, 17]).

2.4.1. Shockley–Read processes The Shockley–Read (SR) mechanism is not an intrinsic and fundamental process as it occurs via levels in the forbidden energy gap. The reported positions of SR centres for both n- and p-type materials range anywhere from near the valence to near the conduction band.

The SR mechanism is probably responsible for lifetimes in low electron concentration n-type $Hg_{1-x}Cd_xTe$. At carrier concentrations less than $10^{15}\,cm^{-3}$, the lifetimes exhibit a broad range of values (0.4–8 μs, $x = 0.20$–0.21, $n \approx 5 \times 10^{14}\,cm^{-3}$) for material prepared by various techniques [17]. The possible factors are centres associated with native defects and residual impurities. Dislocation may also influence the recombination time for dislocation densities $>5 \times 10^5\,cm^{-2}$ [18].

In p-type HgCdTe, the SR mechanism is usually blamed for the reduction of lifetime with decreasing temperature. The steady-state low-temperature photoconductive lifetimes are usually much shorter than the transient lifetimes. The low-temperature lifetimes exhibit very different temperature dependences with a broad range of values over three orders of magnitude, from 1 ns to 1 μs ($p \approx 10^{16}\,cm^{-3}$, $x \approx 0.2$, $T \approx 77$ K, vacancy doping) [17, 19]. This is due to many factors which may affect the measured lifetime including inhomogeneities, inclusions, surface and contact phenomena. Typically, Cu or Au doped materials exhibit lifetimes one order of magnitude larger [13]. It is believed that the increase of lifetime in impurity doped $Hg_{1-x}Cd_xTe$ arises from a reduction of SR centres. This may be due to the low-temperature growth of doped layers or to low-temperature annealing of doped samples.

The origin of the SR centres in p-type material is not clear. These centres do not seem to be the vacancies themselves and thus may be removable [20]. Vacancy doped material with the same carrier concentration, but created under different annealing temperatures may produce different lifetimes. One possible candidate for recombination centres is Hg interstitials [21]. Vacancy doped $Hg_{1-x}Cd_xTe$ exhibits SR recombination centre densities roughly proportional to the vacancy concentration.

Measurements at DSR (formerly Texas Instruments) [22] give lifetimes values for extrinsic material

$$\tau_{ext} = 9 \times 10^9 \frac{p_1 + p}{pN_a}, \tag{11}$$

where

$$p_1 = N_v \exp\left(\frac{q(E_r - E_g)}{kT}\right) \tag{12}$$

and E_r is the SR centre energy relative to the conduction band. Experimentally E_r was found to lie at the intrinsic level for As, Cu and Au dopants, giving $p_1 = n_i$.

For vacancy doped p-type $Hg_{1-x}Cd_xTe$

$$\tau_{vac} = 5 \times 10^9 \frac{n_1}{pN_{vac}}, \tag{13}$$

where

$$n_1 = N_c \exp\left(\frac{qE_r}{kT}\right) \tag{14}$$

E_r is $\approx$30 mV from the conduction band ($x = 0.22$–0.30).

As follows from these expressions, doping with the foreign impurities gives lifetimes which are significantly increased compared with native doping at the same level. Although considerable research effort is still necessary, the SR process does not represent a fundamental limit to the performance of the photodetectors.

2.4.2. Radiative processes For a long time, internal radiative processes have been considered to be the main fundamental limit to detector performance and the performance of practical devices have been compared to that limit. The role of radiative mechanism in the detection of IR radiation has been critically re-examined [23, 24]. Humpreys [23] indicated that most of the photons emitted in photodetectors as a result of radiative decay are immediately re-absorbed, so that the observed radiative lifetime is only a measure of how well photons can escape from the body of the detector. Due to re-absorption the radiative lifetime is highly extended, and dependent on the semiconductor geometry. Therefore, combined recombination–generation processes in one detector are essentially noiseless. In contrast, the recombination act with cognate escape of a photon from the detector, or generation of photons by thermal radiation from outside the active body of the detector are noise producing processes. This may readily happen for a case of detector array, where an element may absorb photons emitted by another detectors or a passive part of the structure [24]. Deposition of the reflective layers (mirrors) on the back and side of the detector may significantly improve optical insulation preventing noisy emission and absorption of thermal photons.

As follows from the above considerations, the internal radiative processes although of fundamental nature, do not limit the ultimate performance of IR detectors.

2.4.3. Auger processes Auger mechanisms dominate generation and recombination processes in high-quality narrow gap semiconductors such as $Hg_{1-x}Cd_xTe$ and InSb at near room temperatures [3]. The band-to-band Auger mechanisms in InSb-like band structure semiconductors are classified in 10 photonless mechanisms. Two of them have the smallest threshold energies ($E_T \approx E_g$) and are denoted as Auger 1 and Auger 7. In some wider band gap materials (e.g. InAs and low x $InAs_{1-x}Sb_x$) in which the split-off band energy Δ is comparable to E_g, the Auger process involving split-off band (AS process) may also play an important role.

The Auger 1 generation is the impact ionization by an electron, generating an electron–hole pair, so this process involves two electrons and one heavy hole, and is dominant in n-type material. An interesting feature is the behaviour of Auger 1 generation and recombination with degenerate n-type doping. Due to the low density of states the Fermi level moves high into the conduction band with n-type doping, so the concentration of minority holes is strongly reduced and the threshold energy required for the Auger transition is increased. This results in suppression of Auger 1 processes in heavy doped n-type material.

Auger 7 generation is the impact generation of electron hole pair by a light hole, involving one heavy hole, one light hole and one electron [25, 26]. This process may dominate in p-type material. Heavy p-type doping does not have such a dramatic effect on the Auger 7 generation and recombination rates due to the much higher density of states. The corresponding Auger recombination mechanisms are inverse processes of the electron–hole recombination with energy transferred to the electron or the hole. Strong temperature and band gap dependence is expected, since lowered temperature and increased band gap strongly reduces the probability of these heat-stimulated transitions.

The net generation rate due to the Auger 1 and Auger 7 processes can be described as [27]

$$G_A - R_A = \frac{n_i^2 - np}{2n_i^2}\left[\frac{n}{(1+an)\tau_{A1}^i} + \frac{p}{\tau_{A7}^i}\right], \tag{15}$$

where $\tau^{\mathrm{i}}_{\mathrm{A1}}$ and $\tau^{\mathrm{i}}_{\mathrm{A7}}$ are the intrinsic Auger 1 and Auger 7 recombination times and n_{i} is the intrinsic concentration. The last equation is valid for a wide range of concentrations, including degeneration, which easily occurs in n-type materials. This is expressed by the finite value of a. According to [27], $a = 5.26 \times 10^{-18}\,\mathrm{cm}^3$. Due to the shape of the valence band, the degeneracy in p-type material occurs only at very high doping levels, which is not achievable in practice.

The Auger 1 lifetime for intrinsic material is equal [28]

$$\tau^{\mathrm{i}}_{\mathrm{A1}} = \frac{h^3\varepsilon_0^2}{2^{3/2}\pi^{1/2}q^4 m_0}\frac{\varepsilon^2(1+\mu)^{1/2}(1+2\mu)\exp[((1+2\mu)/(1+\mu))E_{\mathrm{g}}/kT]}{(m_{\mathrm{e}}^*/m)|F_1F_2|^2(kT/E_{\mathrm{g}})^{3/2}}, \tag{16}$$

where μ is the ratio of the conduction to the heavy-hole valence-band effective mass, ε_{s} is the static-frequency dielectric constant and $|F_1F_2|$ are the overlap integrals of the periodic part of the electron wave functions. The overlap integrals cause the biggest uncertainty in the Auger 1 lifetime. Values ranging from 0.1 to 0.3 have been obtained by various authors. In practice it is taken as a constant equal to anywhere between 0.1 and 0.3 leading to changes by almost an order of magnitude in the lifetime.

The ratio of Auger 7 and Auger 1 intrinsic times

$$\gamma = \frac{\tau^{\mathrm{i}}_{\mathrm{A7}}}{\tau^{\mathrm{i}}_{\mathrm{A1}}} \tag{17}$$

is another term of high uncertainty. According to Casselman and co-workers [25, 26, 29], for $Hg_{1-x}Cd_xTe$ over the range $0.16 \leqslant x \leqslant 0.40$ and $50\,\mathrm{K} \leqslant T \leqslant 300\,\mathrm{K}$, $3 \leqslant \gamma \leqslant 6$. As γ is higher than unity, higher recombination lifetimes are expected in p-type materials when compared to n-type materials of the same doping.

Accurate calculations of the Auger lifetimes have been reported by Beattie and White [30]. The flat valence band model has been used to obtain a simple analytic approximation that requires just two parameters to cover a wide range of temperature and carrier Fermi levels, both degenerate and non-degenerate. Recent direct measurements of carrier recombination show $\gamma = 8$ for $x \approx 0.2$ at 295 K [31].

Kinch [22] delivered a simplified formula for the Auger 1 intrinsic recombination time which is

$$\tau^{\mathrm{i}}_{\mathrm{A1}} = 8.3 \times 10^{-13} E_{\mathrm{g}}^{1/2}\left(\frac{q}{kT}\right)^{3/2}\exp\left(\frac{qE_{\mathrm{g}}}{kT}\right), \tag{18}$$

where E_{g} is in electronvolts.

As equations (16) and (18) show, the Auger generation and recombination rates are strongly dependent on temperature via dependence of carrier concentration and intrinsic time on temperature. Therefore, cooling is a natural and a very effective way to suppress Auger processes.

2.5. α/G ratio and detectivity

The detectivity D^* as the main parameter of detector, characterizes the normalized signal-to-noise performance of the device and is defined as

$$D^* = \frac{R_i(A_0\Delta f)^{1/2}}{I_{\mathrm{n}}}, \tag{19}$$

where R_i is spectral current responsivity

$$R_i = \frac{\lambda\eta}{hc}qg, \tag{20}$$

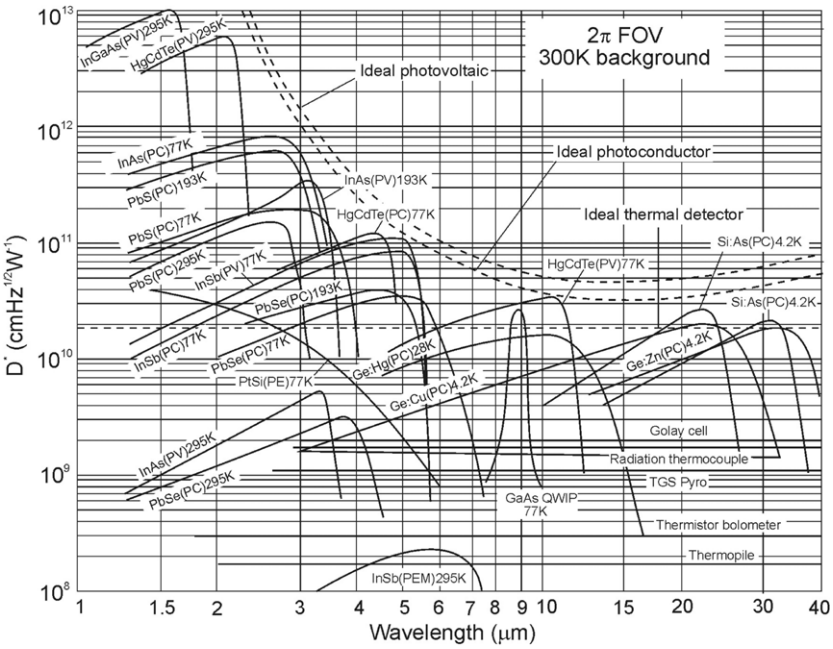


Figure 3. Comparison of the D^* of various commercially available infrared detectors when operated at the indicated temperature. Chopping frequency is 1000 Hz for all detectors except the thermopile (10 Hz), thermocouple (10 Hz), thermistor bolometer (10 Hz), Golay cell (10 Hz) and pyroelectric detector (10 Hz). Each detector is assumed to view a hemispherical surrounding at a temperature of 300 K. Theoretical curves for the background-limited D^* (- - - -) for ideal photovoltaic and photoconductive detectors and thermal detectors are also shown. PC—photoconductive detector, PV—photovoltaic detector, PE—photoemissive detector and PEM—photoelectromagnetic detector.

A_0 is the optical area of detector, Δf is the frequency band, λ is the wavelength, h is Planck's constant, c is the light velocity, q is the electron charge and g is the photoconductive gain.

Spectral detectivity curves for a number of commercially available IR detectors are shown in figure 3. Interest has centred mainly on the wavelengths of the two atmospheric windows 3–5 μm (MWIR region) and 8–14 μm (LWIR region) (atmospheric transmission is the highest in these bands and the emissivity maximum of the objects at $T \approx 300$ K is at the wavelength $\lambda \approx 10$ micron), though in recent years there has been increasing interest in longer wavelengths stimulated by space applications.

Assuming that the current gain for photocurrent and noise current is the same, the current noise due to the generation and recombination processes is [32]

$$I_{\mathrm{n}}^2 = 2(G + R)A_{\mathrm{e}}t\Delta f q^2 g^2, \tag{21}$$

where A_{e} is the electrical area and t is the thickness of detector.

It can be shown that assuming equilibrium conditions ($G = R$) and $A_0/A_{\mathrm{e}} = 1$ [33]

$$D^* = 0.31\frac{\lambda}{hc}k\left(\frac{\alpha}{G}\right)^{1/2}, \tag{22}$$

where $1 \leqslant k \leqslant 2$ is dependent on the contribution of recombination and backside reflection.

The ratio of the absorption coefficient to the thermal generation rate, α/G, is the fundamental figure of merit of any material intended for IR photodetectors. It determines

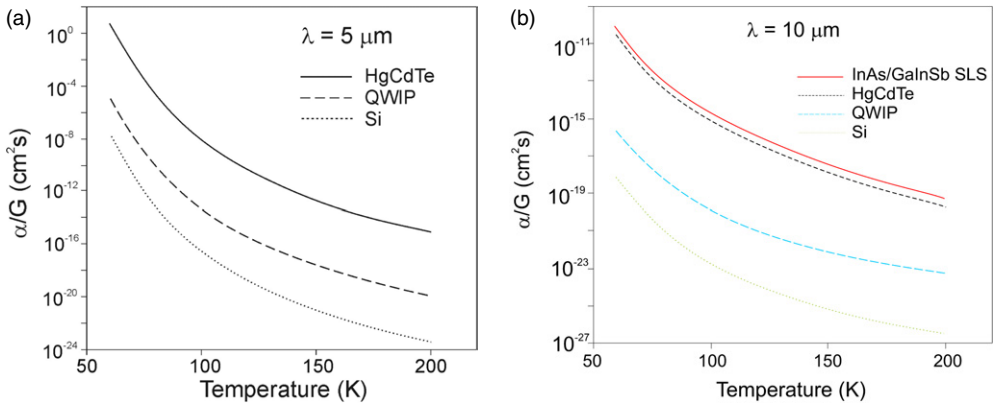


Figure 4. α/G ratio versus temperature for MWIR—$\lambda = 5\,\mu$m (*a*) and LWIR—$\lambda = 10\,\mu$m (*b*) photon detectors.

directly the detectivity limits of the devices. Any potential material should be compared on this basis.

The α/G ratio versus temperature for different types of tunable materials with hypothetical energy gap equal to 0.25 eV ($\lambda = 5\,\mu$m) and 0.124 eV ($\lambda = 10\,\mu$m) is shown in figure 4. Procedures used in calculations of α/G for different material systems are given in [34]. Estimation of α/G ratio for InAs/GaInSb strained layer superlattices (SLSs) are based on several theoretical papers published previously [35–39]. Early calculations showed that a LWIR type-II InAs/GaInSb SLS should have an absorption coefficient comparable to a HgCdTe alloy with the same cutoff wavelength [35]. Figure 4(*b*) predicts that type-II SLs is the most efficient detector of IR radiation in long wavelength region. It is an even better material than HgCdTe; it is characterized by a high absorption coefficient and relatively low thermal generation rate. However hitherto, this theoretical prediction has been not confirmed by experimental data (see section 8.4). We can also notice that AlGaAs/GaAs quantum well IR photoconductor (QWIP) is a better material than extrinsic silicon.

The total generation rate is a sum of the optical and thermal generation

$$G = G_{\mathrm{th}} + G_{\mathrm{op}}. \tag{23}$$

The optical generation may be due to the signal or background radiation. For IR detectors, background radiation is usually higher compared to the signal radiation. If the thermal generation is reduced much below the background level, the performance of the device is determined by the background radiation (BLIP conditions for background limited infrared photodetector). This condition can be described as

$$\frac{\eta \Phi_{\mathrm{B}} \tau}{t} > n_{\mathrm{th}}, \tag{24}$$

where n_{th} is the density of thermal carriers at temperature T, τ is the carrier lifetime and Φ_{B} is the total background photon flux density (unit $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$) reaching the detector. Rearranging we have for the BLIP requirements

$$G_{\mathrm{op}} = \frac{\eta \Phi_{\mathrm{B}}}{t} > \frac{n_{\mathrm{th}}}{\tau} = G_{\mathrm{th}}, \tag{25}$$

i.e. the photon generation rate per unit volume needs to be greater than the thermal generation rate per unit volume. The carriers can be either majority or minority in nature.

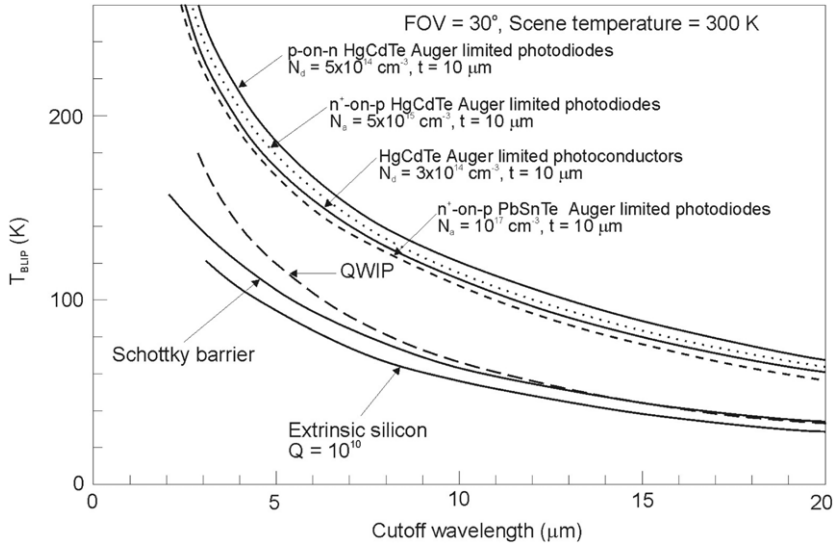


Figure 5. Estimation of the temperature required for background limited operation of different types of photon detectors. In the calculations FOV $= 30^\circ$ and $T_B = 300$ K are assumed.

The background limited detectivity, or the so-called 'photovoltaic' BLIP detectivity, is given by [40]

$$D^*_{\mathrm{BLIP}} = \frac{\lambda}{hc}\left(\frac{\eta}{2\Phi_{\mathrm{B}}}\right)^{1/2}. \tag{26}$$

D^*_{BLIP} for photoconductors is $\sqrt{2}$ times lower than that for photodiodes. This is attributable to the recombination process in photoconductors, which is uncorrelated with the generation process, which contributes to the detector noise. The background photon flux density received by the detector depends on its angular field of view (FOV) of the background and on its ability to respond to the wavelengths contained in this source.

BLIP temperature is defined as the device operating at a temperature at which the dark current equals the background photocurrent, given a FOV and a background temperature. In figure 5 plots of the calculated temperature required for background limited operation in 30° FOV are shown as a function of cutoff wavelength. We can see that the operating temperature of 'bulk' intrinsic IR detectors (HgCdTe and PbSnTe) is higher than for other types of photon detectors. HgCdTe detectors with background limited performance operate with thermoelectric coolers in the MWIR range, but the LWIR detectors ($8 \leqslant \lambda_c \leqslant 12\,\mu$m) operate at $\approx$100 K. HgCdTe photodiodes exhibit a higher operating temperature compared with extrinsic detectors, silicide Schottky barriers and QWIPs. However, the cooling requirements for QWIPs with cutoff wavelengths below $10\,\mu$m are less stringent in comparison with extrinsic detectors and Schottky barrier devices.

3. Photodiodes—principle of operation and figure of merit

The most common example of a photovoltaic detector is the abrupt p–n junction prepared in the semiconductor, which is often referred to simply as a photodiode. The operation of the p–n junction photodiode is illustrated in figure 6. Photons with energy greater than the energy gap,

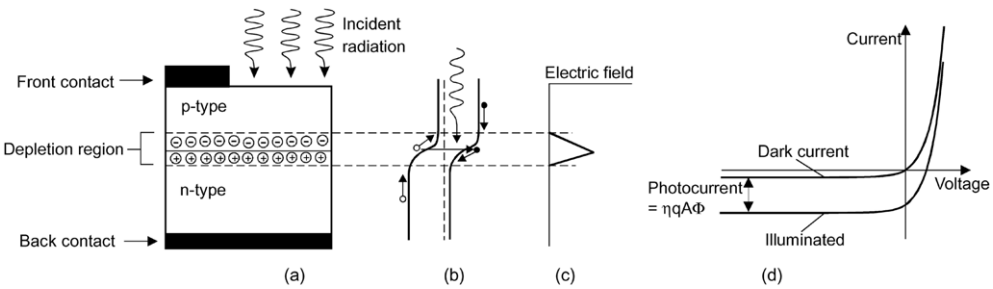


Figure 6. p–n junction photodiode: (*a*) structure of abrupt junction, (*b*) energy band diagram, (*c*) electric field and (*d*) I–V characteristics.

incident on the front surface of the device, create electron–hole pairs in the material on both sides of the junction. By diffusion, the electrons and holes generated within a diffusion length from the junction reach the space–charge region. Then electron–hole pairs are separated by the strong electric field; minority carriers are readily accelerated to become majority carriers on the other side. In this way a photocurrent is generated which shifts the current–voltage (I–V) characteristic in the direction of negative or reverse current. The total current density in the p–n junction is usually written as

$$J(V, \Phi) = J_{\mathrm{d}}(V) - J_{\mathrm{ph}}(\Phi), \tag{27}$$

where the dark current density, J_{d}, depends only on the bias voltage, V and the photocurrent depends only on the photon flux density, Φ.

Generally, the current gain in a simple photovoltaic detector (e.g. not an avalanche of photodiodes) is equal to 1, and the magnitude of photocurrent is equal

$$I_{\mathrm{ph}} = \eta q A \Phi. \tag{28}$$

The dark current and photocurrent are linearly independent (which occurs even when these currents are significant).

If the p–n diode is open-circuited, the accumulation of electrons and holes on the two sides of the junction produces an open-circuit voltage (figure 6(*d*)). If a load is connected to the diode, a current will be conducted in the circuit. The maximum current is realized when an electrical short is placed across the diode terminals and this is called the short-circuit current.

The open-circuit voltage can be obtained by multiplying the short-circuit current by the incremental diode resistance $R = (\partial I/\partial V)^{-1}$ at $V = V_{\mathrm{b}}$:

$$V_{\mathrm{ph}} = \eta q A \Phi R, \tag{29}$$

where V_{b} is the bias voltage and $I = f(V)$ is the I–V characteristic of the diode.

In many direct applications the photodiode is operated at zero-bias voltage:

$$R_0 = \left(\frac{\partial I}{\partial V}\right)^{-1}_{|V_{\mathrm{b}}=0}. \tag{30}$$

A frequently encountered figure of merit for IR photodiode is the $R_0 A$ product

$$R_0 A = \left(\frac{\partial J}{\partial V}\right)^{-1}_{|V_{\mathrm{b}}=0}, \tag{31}$$

where $J = I/A$ is the current density.

In the detection of radiation, the photodiode is operated at any point of the I–V characteristic. High reverse bias may or may not shorten charge collection time, but it generally

reduces capacitance. Both result in a faster response. On the other hand, increased reverse bias causes increased noise, so that a tradeoff exists between speed and sensitivity.

The intrinsic noise mechanism of a photodiode is shot noise in the current passing through the diode. It is generally accepted that the noise in an ideal diode is given by

$$I_n^2 = [2q(I_d + 2I_s) + 4kT(G_j - G_0)]\Delta f, \tag{32}$$

where $I_d = I_s[\exp(qV/kT) - 1]$ is the dark current of the ideal diode (diffusion current), G_j is the conductance of the junction and G_0 is the low-frequency value of G_j. In the low-frequency region, the second term on the right-hand side is zero.

Photodiodes are typically operated at zero bias to minimize the heat load and for zero $1/f$ noise. In this case

$$I_n^2 = 2(2G + \eta\Phi_B)q^2 A\Delta f. \tag{33}$$

Frequently, the performance of photodiodes is discussed in terms of zero bias resistance-area product. As follows for the ideal diode I–V plot (see equation (31))

$$R_0 A = \frac{kT}{qJ_s} = \frac{kT}{q^2 Gt}, \tag{34}$$

since $J_s = qGt$, where t is the thickness of the photodiode's active region.

Taking into account the Auger 7 mechanism in extrinsic p-type region of n^+-on-p photodiode, we receive

$$(R_0 A)_{A7} = \frac{2kT\tau_{A7}^i}{q^2 N_a t} \tag{35}$$

and the same equation for P-on-n photodiode (with dominant contribution of Auger 1 mechanism in n-type region)

$$(R_0 A)_{A1} = \frac{2kT\tau_{A1}^i}{q^2 N_d t}, \tag{36}$$

where N_a and N_d are the acceptor and donor concentrations in the base regions, respectively.

As equations (35) and (36) show, the $R_0 A$ product can be decreased by reduction of the thickness of the base layer. Since $\gamma = \tau_{A7}^i/\tau_{A1}^i > 1$, a higher $R_0 A$ value can be achieved in p-type base devices compared to that of n-type devices of the same doping level. Detailed analysis shows that the absolute maximum of $R_0 A$ is achievable with base layer doping producing $p = \gamma^{1/2} n_i$, which corresponds to the minimum of thermal generation. The required p-type doping is difficult to achieve in practice for low temperature HgCdTe photodiodes (the control of hole concentration below $5 \times 10^{15}\,\mathrm{cm}^{-3}$ level is difficult) and the p-type material suffer from some non-fundamental limitations, such as contacts, surface and SR processes. These are the reasons why the low temperature detectors are typically produced from the lightly doped n-type materials.

The zero bias detectivity can be expressed as

$$D^* = \frac{\lambda\eta q}{hc}\frac{(R_0 A)^{1/2}}{(4kT + 2q^2\eta\Phi_B R_0 A)^{1/2}}. \tag{37}$$

For the best performance under the given operating conditions (wavelength, temperature) the value of $\eta(R_0 A)^{1/2}$ should be maximized. The $\eta(R_0 A)^{1/2}$ is a photodiode figure of merit that determines the performance of a photodiode.

Figure 7 illustrates the detectivity that can be achieved for p-on-n HgCdTe photodiodes for four selected wavelength regions. At low temperatures, the detector thermal noise is negligible, and detectivity is limited by detector noise due to fluctuations in the arrival rate of photons from

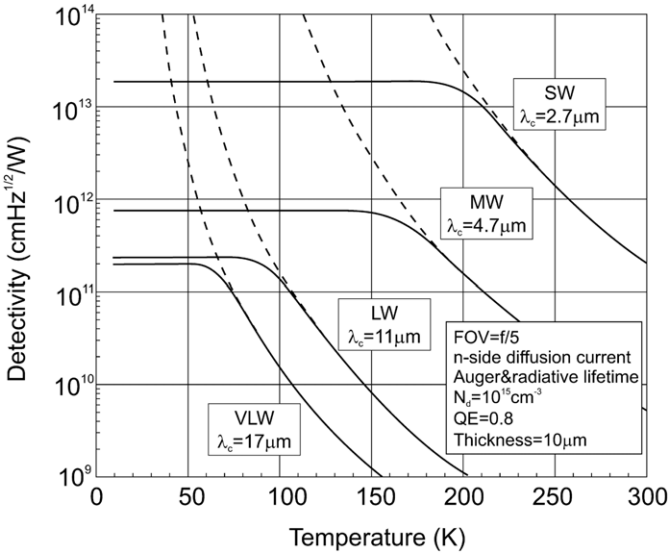


Figure 7. Calculated detectivity for p-on-n HgCdTe photodiodes for four important wavelength regions, plotted versus operating temperature (after [2]).

room-temperature background radiation (BLIP operation). As detector temperature increases, the detector thermal noise increases exponentially, and usually overcomes the background noise, causing the detectivity to decrease exponentially for further increases in temperature. Detector thermal noise is proportional to the thermal generation rate (see equation (21)),which is inversely proportional to the carrier lifetime (see equation (25)).

Hitherto, we have assumed that the photodiode's dark current was limited by diffusion. In general, the dark current mechanisms can be classified into two groups:

- inherent or fundamental mechanisms, which depend only on the inherent material properties:
  - diffusion current due to Auger or radiative recombination in the n-region or p-region,
  - band-to-band tunnelling current,
- defect-related mechanisms which require surface or 'bulk' defects located within depletion region or within a diffusion length of either side of the depletion region:
  - diffusion current due to SR recombination in the n-region or p-region
  - generation–recombination within the depletion region,
  - trap-assisted tunnelling,
  - surface generation current from surface states.

The junction mechanisms in HgCdTe photodiodes have been reviewed by many authors, e.g. [3, 41–43]. We will be concerned with the current contribution of high-quality photodiodes with high R_0A products limited by the following:

- generation–recombination within the depletion region,
- tunnelling through the depletion region,
- surface effects.

The first two mechanisms are schematically illustrated in figure 8.

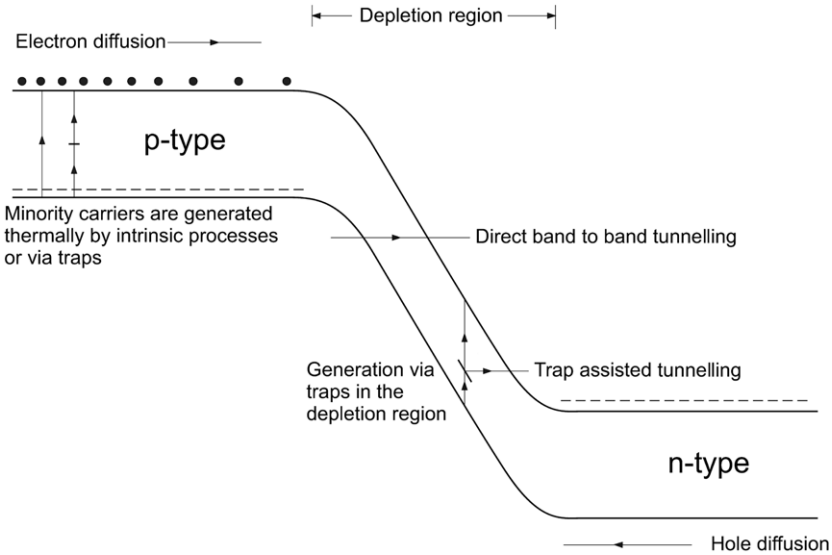


Figure 8. Schematic representation of some of the mechanisms by which dark current is generated in a reverse biased p–n junction.

The generation–recombination current of the depletion region, for the simple case of uniformly distributed bulk and surface g–r centres, can be described as

$$J_{\mathrm{GR}} = \frac{q w n_{\mathrm{i}}}{\tau_0}. \tag{38}$$

In this equation, τ_0 is an effective carrier lifetime in depletion region and w is the depletion region width at zero bias voltage.

The space–charge region generation–recombination current varies with temperature as n_{i}, i.e. less rapidly than diffusion current which varies as n_{i}^2.

The zero bias resistance area product for the case of uniformly distributed bulk and surface g–r centres is given by

$$\left(\frac{1}{R_0 A}\right)_{\mathrm{GR}} = \frac{e n_{\mathrm{i}} w}{V_{\mathrm{b}}}\left(\frac{1}{\tau_0} + \frac{S_0 P}{A}\right), \tag{39}$$

where V_{b} is the built-in voltage of the p–n junction. The second term in parentheses is for g–r centres localized at the surface and the parameter S_0 is referred to as the g–r surface recombination velocity, which is proportional to the density of g–r defects. P is the junction perimeter. Note that the bulk and surface components of the g–r $R_0 A$ product have different dependencies of junction geometry.

The third type of dark current component that can exist is a tunnelling current caused by electrons directly tunnelling across the junction from the valence band to the conduction band (direct tunnelling) or by electrons indirectly tunnelling across the junction by way of intermediate trap sites in the junction region (indirect tunnelling or trap-assisted tunnelling—see figure 8).

The tunnelling current (and $R_0 A$ product) is critically dependent on doping concentration. Figure 9 shows the dependence of the $R_0 A$ product components on the dopant concentrations for one-sided abrupt HgCdTe, PbSnTe and PbSnSe photodiodes at 77 K ($E_{\mathrm{g}} = 0.1$ eV). To produce high $R_0 A$ products for HgCdTe and lead salt photodiodes, the doping concentration of

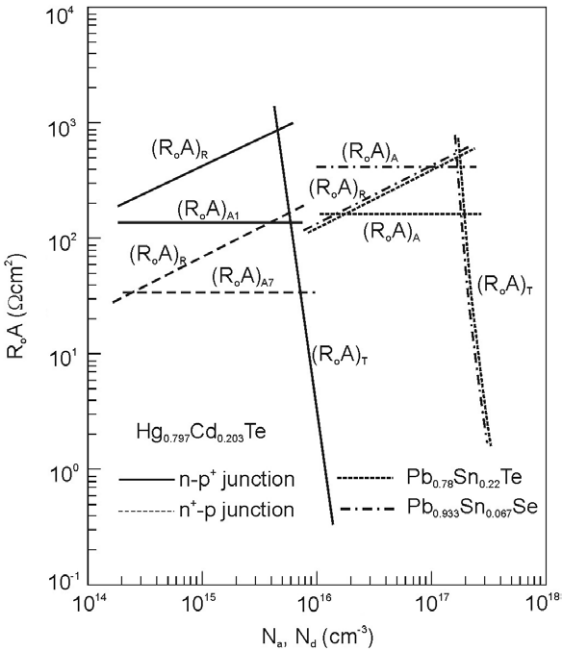


Figure 9. The dependence of the R_0A product components (due to diffusion current with radiative—$(R_0A)_R$ and Auger—$(R_0A)_A$ processes, and due to band-to-band tunnelling—$(R_0A)_T$) on the dopant concentrations for the one-sided abrupt junctions produced in $Hg_{0.797}Cd_{0.203}Te$, $Pb_{0.78}Sn_{0.22}Te$ and $Pb_{0.933}Sn_{0.067}Se$ (after [44]).

10^{16} cm^{-3} and 10^{17} cm^{-3} (or less) are required, respectively. The maximum available doping levels due to onset of tunnelling are more than an order of magnitude higher with IV–VI than with HgCdTe photodiodes [44]. This is due to their high permittivities ε_s because tunnelling contribution of the R_0A product contains $\exp[\mathrm{const}(m^*\varepsilon_s/N)^{1/2}E_g]$ factor.

In a real p–n junction additional dark current related to the surface occurs. The surface of actual devices is passivated in order to stabilize the surface against chemical and heat-induced changes as well as to control surface recombination, leakage and related noise. An ideal surface would be electrically neutral and would have a very low density of surface states.

The kinetics of generation–recombination through fast surface states is identical to that through bulk SR centres. The current in a surface channel is given by (see equation (38))

$$I_{\mathrm{GRS}} = \frac{qn_i w_c A_c}{\tau_0}, \tag{40}$$

where w_c is the channel width and A_c is the channel area.

Apart from generation–recombination processes occurring at the surface and within surface channels there are other surface-related current mechanisms, termed surface leakage, with ohmic or breakdown-like I–V characteristics. They are nearly temperature independent.

4. Historical perspective

The first report of the synthesis of the semi-metal HgTe and wide band gap semiconductor CdTe to form a HgCdTe ternary alloy system [1] was published in 1959 by the Royal Radar Establishment in Malvern, England. This article reported both photoconductive and

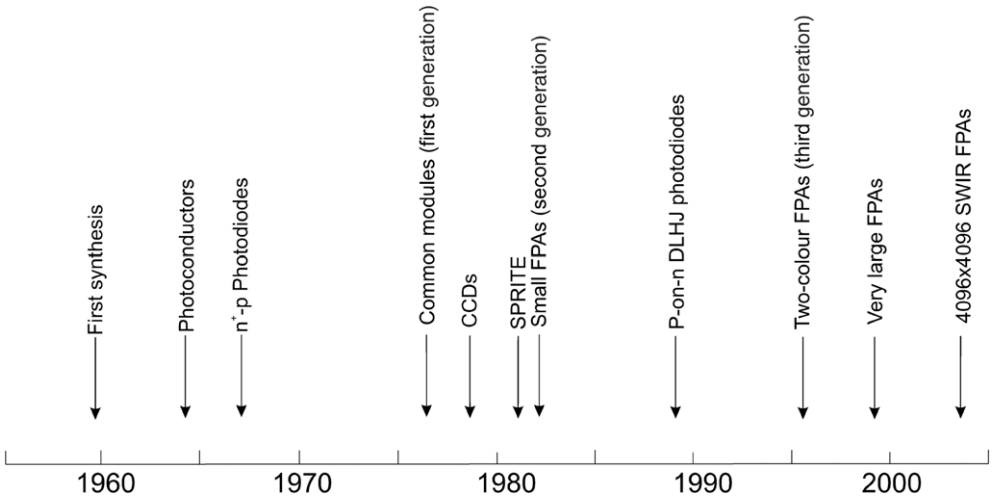


Figure 10. History of the development of HgCdTe detectors.

photovoltaic responses at wavelengths extending out to 12 μm, and made the understated observation that this material showed promise for intrinsic IR detectors. At that time the importance of the 8–12 μm atmospheric transmission window was well known for thermal imaging, which enables night vision by imaging the emitted IR radiation from the scene. Since 1954, the Cu-doped extrinsic photoconductive detector was known [45], but its spectral response extended to 30 μm (far longer than required for the 8–12 μm window) and to achieve background-limited performance the Ge : Cu detector was necessary to cool down to liquid helium temperature. In 1962, it was discovered that the Hg acceptor level in Ge has an activation energy of about 0.1 eV [46] and detector arrays were soon made from this material; however, the Ge : Hg detectors were cooled to 30 K to achieve maximum sensitivity. It was also clear from theory that intrinsic HgCdTe detectors (where the optical transitions were direct transitions between the valence band and the conduction band) could achieve the same sensitivity at much higher operating temperature (as high as 77 K). Early recognition of the significance of this fact led to intensive development of HgCdTe detectors in a number of countries including England, France, Germany, Poland, the former Soviet Union and the United States [47]. However, little has been written about the early development years; e.g. the existence of work going on in the United States was classified until the late 1960s.

Figure 10 gives approximate dates of significant development efforts for HgCdTe IR detectors; and figure 11 gives additional insight in time line of the evolution of detectors and key developments in process technology. Photoconductive devices had been built in the US as early as 1964, at Texas Instruments, after the development of a modified Bridgman crystal growth technique. The first report of a junction intentionally formed to make an HgCdTe photodiode was by Verie and Granger [49], who used Hg in-diffusion into p-type material doped with Hg vacancies. The first important application of HgCdTe photodiodes was as high-speed detectors for CO_2 laser radiation [50]. The French pavilion at the 1967 Montreal Expo illustrated a CO_2 laser system with a HgCdTe photodiode. However, the high performance MWIR and LWIR linear arrays developed and manufactured in the 1970s were n-type photoconductors used in the first generation scanning systems. In 1969 Bartlett *et al* [51] reported background limited performance of photoconductors operated at 77 K in the LWIR spectral region. The advantages in material preparation and detector technology have led to devices approaching

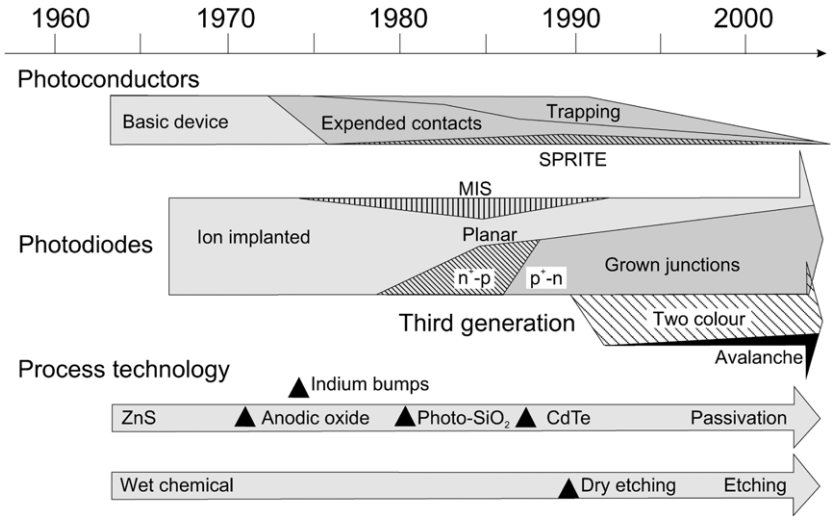


Figure 11. A time line of the evolution of HgCdTe IR detectors and key developments in process technology which made them possible (after [48]).

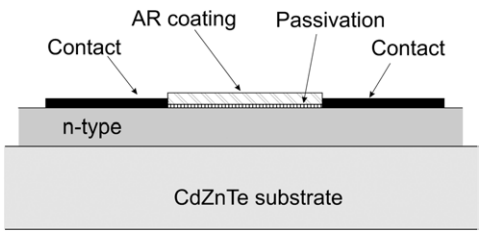


Figure 12. Cross section of a basic HgCdTe photoconductor.

the theoretical limits of responsivity and detectivity over wide ranges of temperature and background [52].

The basic photoconductive device structure is illustrated in figure 12, where metal electrodes are applied to pure n-type material thinned to approximately 10 μm. Typical photoconductors are passivated with anodic oxide and anti-reflection coated with zinc sulphide. The photoconductors required far more simple materials growth and device processing technologies than p–n junctions.

A novel variation of the standard photoconductive device, the SPRITE detector (the acronym of Signal PRocessing In The Element), was invented in England [53]. A family of thermal imaging systems has utilized this device; however, a decline in its usage is observed now. The SPRITE detector provides signal averaging of a scanned image spot which is accomplished by synchronization between the drift velocity of minority carriers along the length of photoconductive bar of material and the scan velocity of the imaging system (see figure 13). Then the image signal builds up a bundle of minority charges which is collected at the end of the photoconductive bar, effectively integrating the signal for a significant length of time and thereby improving the signal-to-noise ratio.

The SPRITE detectors are fabricated from lightly doped ($\approx 5 \times 10^{14}$ cm^{-3}) $Hg_{1-x}Cd_xTe$. Both bulk material and epilayers are being used. Single and 2, 4, 8, 16 and 24 element arrays have been demonstrated; the 8-element arrays are the most commonly used (figure 14).

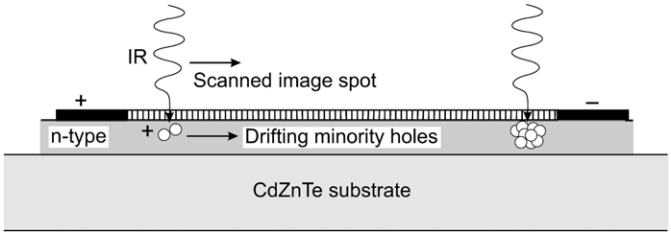


Figure 13. Cross section of a SPRITE photoconductive detector.

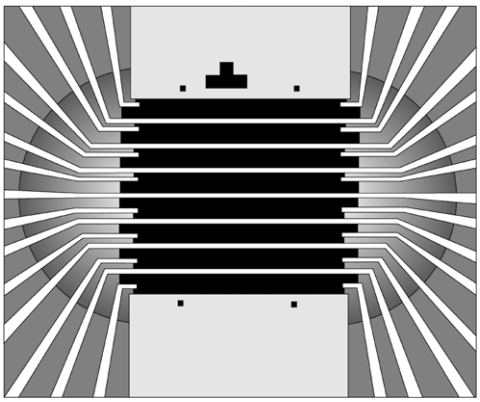


Figure 14. Schematic of eight-element SPRITE array with bifurcated readout zones (after [54]).

The simplest scanning linear FPA consists of a row of detectors (see figure 15). An image is generated by scanning the scene across the strip using, as a rule, a mechanical scanner. These types of arrays can provide enhanced sensitivity and gain in camera weight. In a general sense, the signal-to-noise ratio of a sensor will improve with the square root of the number of detector elements in an array—to the extent that they can collect proportionally more signal from the scene. The scanning system, which does not include multiplexing functions in the focal plane, belongs to the first generation systems. A typical example of this kind of detector is a linear photoconductive array in which an electrical contact for each element of a multi-element array is brought off the cryogenically-cooled focal plane to the outside, where there is one electronic channel at ambient temperature for each detector element. The US common module HgCdTe arrays employ 60, 120 or 180 photoconductive elements depending on the application (see figure 16).

After the invention of charge coupled devices (CCDs) by Boyle and Smith [55] the idea of an all-solid-state electronically scanned two-dimensional (2D) IR detector array caused attention to be turned to HgCdTe photodiodes. These include p–n junctions, heterojunctions and MIS photo-capacitors. Each of these different types of devices has certain advantages for IR detection, depending on the particular application. More interest has been focused on the first two structures, so further considerations are restricted to p–n junctions and heterostructures. Photodiodes with their very low power dissipation, inherently high impedance, negligible $1/f$ noise and easy multiplexing on focal plane silicon chip can be assembled in 2D arrays containing a very large number of elements limited only by existing technologies. They can be reverse-biased for even higher impedance, and can, therefore, match electrically with compact

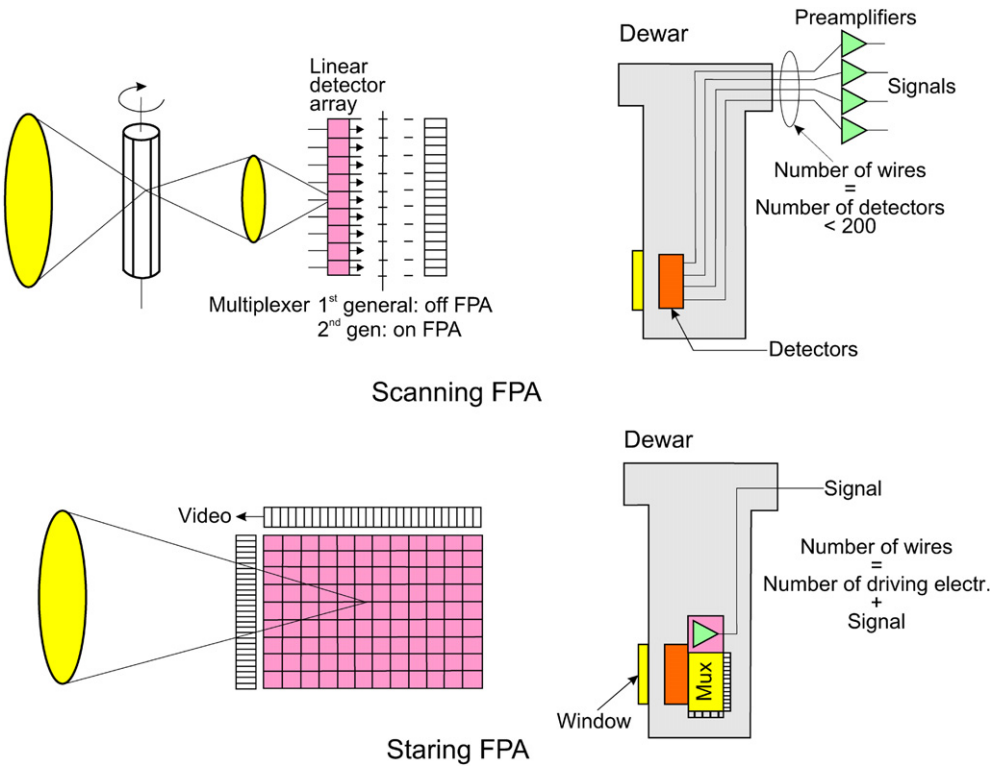


Figure 15. Focal plane arrays: scanning and staring systems.

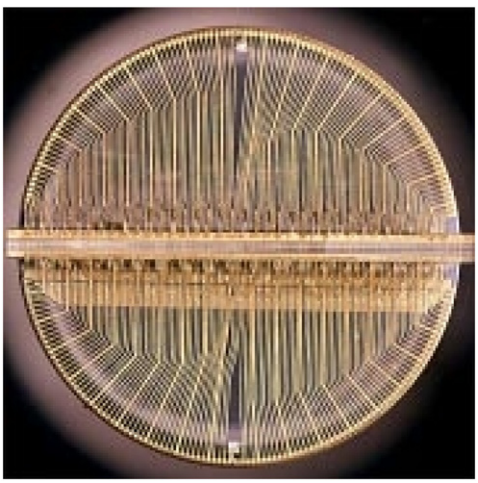


Figure 16. 180 element common module linear array of HgCdTe photoconductive detectors.

low-noise silicon readout preamplifier circuits. The response of photodiodes remains linear to significantly higher photon flux levels than that of photoconductors (because of higher doping levels in the photodiode absorber layer and because the photogenerated carriers are collected rapidly by the junction). At the end of 1970s the emphasis was directed towards

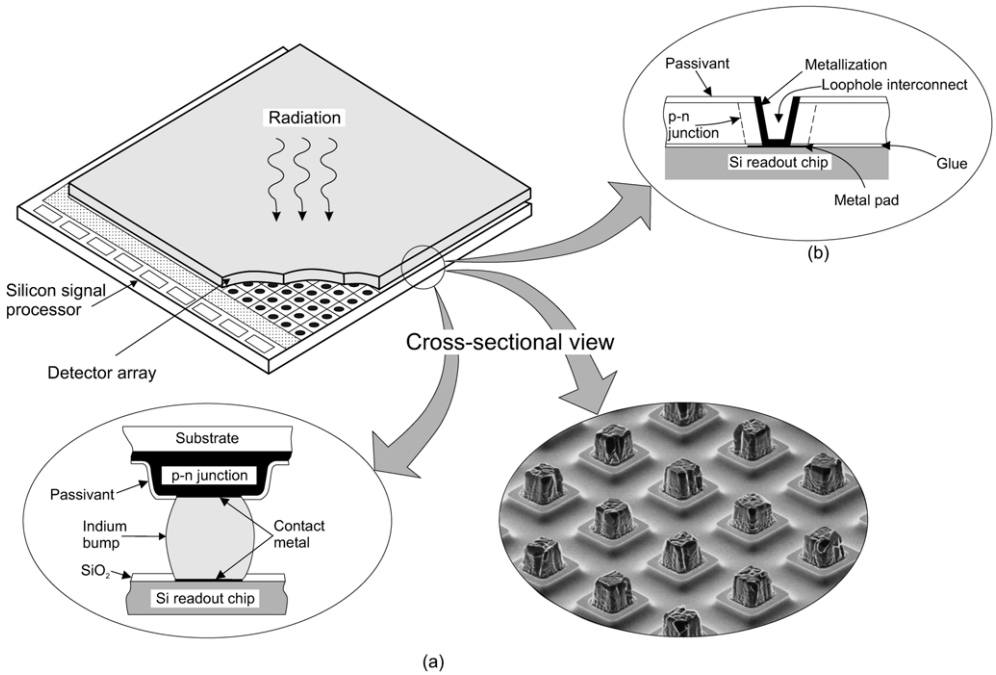


Figure 17. Hybrid IR FPA with independently optimized signal detection and readout: (*a*) indium bump technique and (*b*) loophole technique.

large photovoltaic HgCdTe arrays in the MW and LW spectral bands for thermal imaging. Recent efforts have been extended to short wavelengths, e.g. for starlight imaging in the SW range, as well as to VLWIR space borne remote sensing beyond 15 μm.

The second generation systems (full-framing systems) have typically three orders of magnitude more elements ($>10^6$) on the focal plane than first generation systems and the detector elements are configured in a 2D array. These starring arrays are scanned electronically by circuits integrated with the arrays. These readout integrated circuits (ROICs) include, e.g. pixel deselecting, anti-blooming on each pixel, subframe imaging, output preamplifiers and some other functions. The second generation HgCdTe devices are 2D arrays of photodiodes. This technology had begun in the late 1970s and took the next decade to reach volume production. The first demonstrated hybrid architecture [56] in the mid-1970s (see figure 17), indium bump bonding of readout electronics provides for multiplexing the signals from thousands of pixels onto a few output lines, greatly simplifying the interface between the vacuum-enclosed cryogenic sensor and the system electronics. The detector material and multiplexer are optimized independently. Other advantages of the hybrid FPAs are near 100% fill factors and increased signal-processing area on the multiplexer chip.

Hybrid FPAs detectors and multiplexers are also fabricated using loophole interconnection [57]. In this case, the detector and the multiplexer chips are glued together to form a single chip before detector fabrication. The photovoltaic detector is formed by ion implantation and loopholes are drilled by ion milling and electrical interconnection between each detector and its corresponding input circuit is made through a small hole formed in each detector. The junctions are connected down to the silicon circuit by cutting the fine, few microns in diameter, holes through the junctions by ion milling and then backfilling the holes

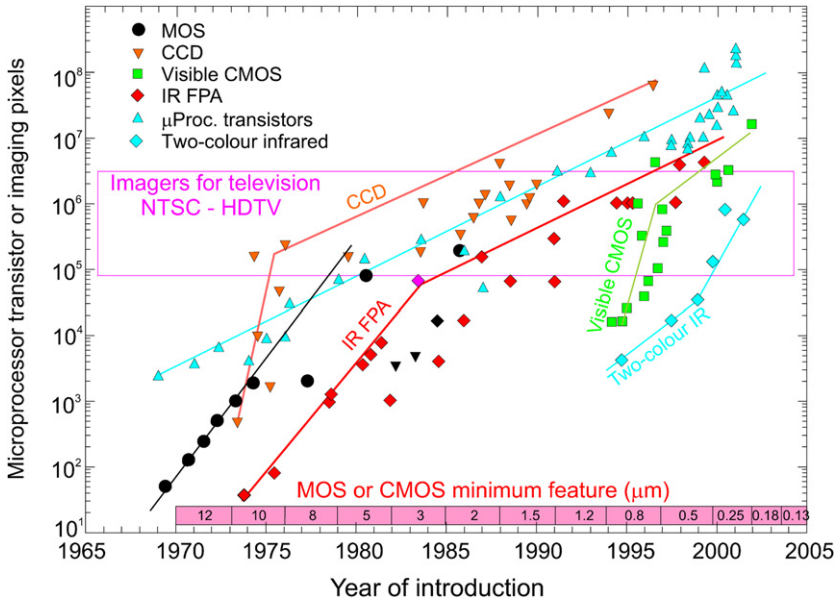


Figure 18. Imaging array formats compared with the complexity of microprocessor technology as indicated by transistor count. The timeline design rule of MOS/CMOS features is shown at the bottom (after [58]).

with metallization. The loophole interconnection technology offers more stable mechanical and thermal features than flip-chip hybrid architecture.

Intermediary systems are also fabricated with multiplexed scanned photodetector linear arrays in use and with, as a rule, time delay and integration (TDI) functions. Typical examples of these systems are HgCdTe multilinear 288×4 arrays fabricated by Sofradir both for 3–5 μm and 8–10.5 μm bands with signal processing in the focal plane.

CMOS multiplexers are the best choice for performing the integration and signal processing for 2D arrays. The advantages of CMOS are that the existing foundries, which fabricate specific integrated circuits, can be readily used by adapting their design rules. Design rules of 0.18 μm are in production with pre-production runs of 0.13 μm design rules. As a result of such fine design rules, more functionality has been put into the unit cells of IR and visible multiplexers and smaller unit cells, leading to large array sizes. Figure 18 shows the timelines for minimum circuit features and the resulting CCD, IR FPA and CMOS visible imager sizes with respect to imaging pixels. Along the horizontal axis is also a scale depicting the general availability of various MOS and CMOS processes. The ongoing migration to even finer lithographies will thus enable the rapid development of CMOS-based imagers having even higher resolution, better image quality, higher levels of integration and lower overall imaging system cost than CCD-based solutions. At present, CMOS with minimum features of $\leqslant 0.5$ μm makes possible also monolithic visible CMOS imagers because the denser photolithography allows for low-noise signal extraction and high performance detection with the optical fill factor within each pixel. The silicon wafer production infrastructure, which has put personal computers into many homes, now allows CMOS-based imaging in consumer products such as video and digital still cameras.

The most commonly used technique for fabricating HgCdTe photodiodes is to make simple n^+-on-p homojunctions by ion implantation into p-type undoped crystals [3].

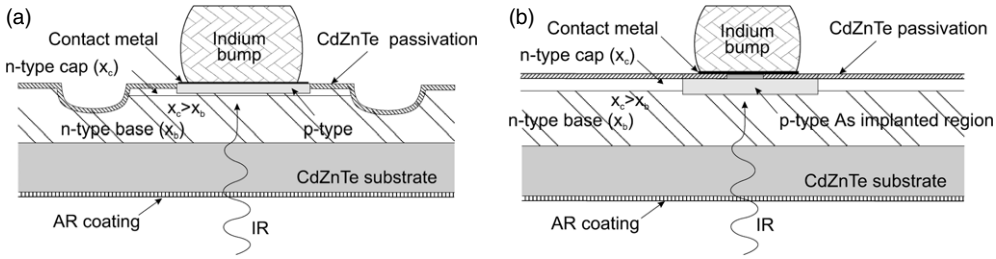


Figure 19. Cross section of p-on-n HgCdTe DLHJ photodiodes: (*a*) mesa structure and (*b*) planar structure. The active n-type region is sandwiched between CdZnTe substrate and high-doped, wider-gap regions.

In 1985, Rogalski and Larkowski [44] indicated that due to the lower minority-carrier diffusion length (lower mobility of holes) in the n-type region of p^+-on-n junctions with thick n-type active region, the diffusion-limited R_0A product of such junctions is larger than for n^+-on-p ones (see figure 9). These theoretical predictions were next confirmed by experimental results obtained for p-on-n HgCdTe junctions, especially for double-layer heterojunction (DLHJ) photodiodes. In such diode (see figure 19) the base n-type layer is sandwiched between CdZnTe substrate and wider-gap regions. Due to backside illumination (through CdZnTe substrate) and internal electric fields (which are 'blocking' for minority carriers), influence of surface recombination on the photodiode performance is eliminated. The influence of surface recombination can be also prevented by the use of suitable passivation. Both optical and thermal generations are suppressed in the p-region due to wide gap.

The MIS photocapacitor is usually formed on an n-type absorber layer with a thin semi-transparent metal film as a gate electrode. The insulator of choice is a thin native oxide. The only motivation for developing HgCdTe MIS detectors was the allure of realizing a monolithic IR CCD, with both the detection and multiplexing taking place in the same material. However, because of the non-equilibrium operation of the MIS detector (usually a bias voltage pulse of several volts is applied across the capacitor to drive the surface into deep depletion), much larger electric fields are set up in the depletion region of the MIS device than in the p–n junction, resulting in defect-related tunnelling current that is orders of magnitude larger than the fundamental dark current. The MIS detector required much higher quality material than photodiode, which has still not been achieved. For this reason, the development of the HgCdTe MIS detector was abandoned around 1987 [59, 60].

In the last decade of the twentieth century, a third generation of HgCdTe detectors emerged from the tremendous impetus given to detector developments (see section 7). This generation of detectors has emerged from technological achievements in the growth of heterostructure devices used in the production of second-generation systems.

5. Impact of epitaxial growth on development of HgCdTe detectors

The timeline for the evolution of growth technologies is illustrated in figure 20. Historically, crystal growth of HgCdTe has been a major problem mainly because a relatively high Hg pressure is present during growth, which makes it difficult to control the stoichiometry and composition of the grown material. The liquidus and solidus lines in the pseudobinary HgTe–CdTe system are shown in figure 21. The wide separation between the liquidus and solidus, leading to marked segregation between CdTe and HgTe, was instrumental in the development of all the bulk growth techniques to this system. In addition to solidus–liquidus separation,

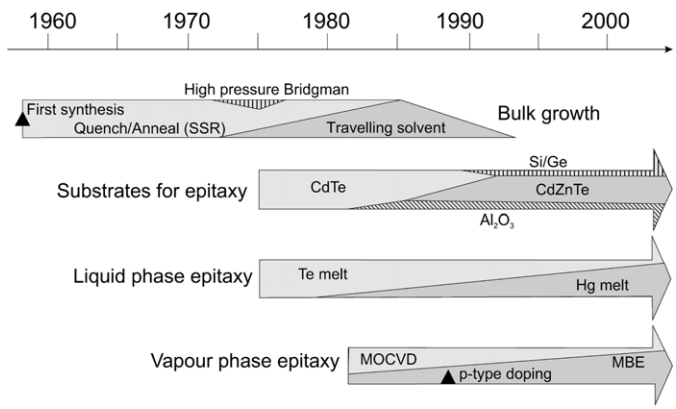


Figure 20. Evolution of HgCdTe crystal growth technology from 1958 to the present (after [48]).

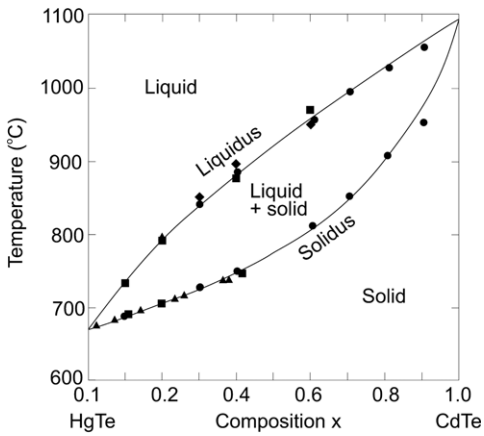


Figure 21. Liquidus and solidus lines in the HgTe–CdTe pseudobinary system (after [61]).

high Hg partial pressures are also influential both during growth and post-growth heat treatments.

Early experiments and a significant fraction of early production were done using a quench–anneal or solid-state recrystallization process. In this method the charge of a required composition was synthesized, melted and quenched. Then, the fine dendritic mass (highly polycrystalline solid) obtained in a process was annealed below the liquidus temperature for a few weeks to recrystallize and homogenize the crystals. The material usually requires low-temperature annealing for adjusting the concentration of native defects. The crystals can be also uniformly doped by the introduction of dopants to the charge.

Bridgman growth was attempted for several years around the mid-1970s. At the same time, solvent growth methods from Te-rich melts were initiated to reduce the growth temperature. One successful implementation was the travelling heater method up to 5 cm diameter [62]. The perfect quality of crystals grown by this method is accompanied by a low growth rate.

The bulk HgCdTe crystals were initially used for many types of IR photodetectors. At present they are still used for some IR applications such as n-type single element photoconductors, SPRITE detectors and linear arrays. Bulk growth produced thin rods,

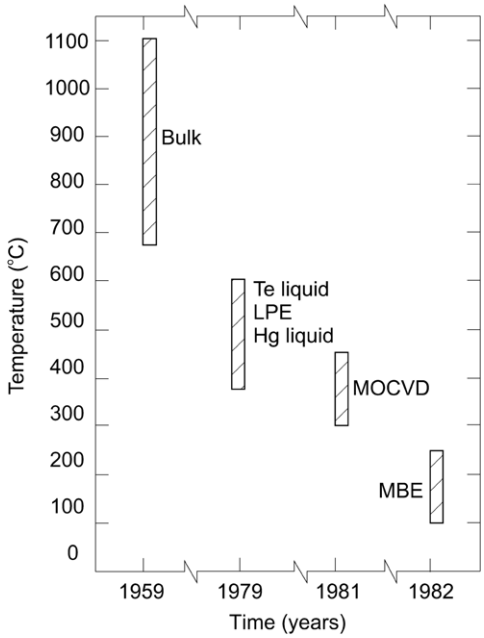


Figure 22. Temperature ranges for growth of HgCdTe by various growth techniques versus date of first reported attempt (after [63]).

generally to 15 mm in diameter, about 20 cm in length and with non-uniform distribution of composition. Large 2D arrays could not be realized with bulk crystals. Another drawback to bulk material was the need to thin the bulk wafers, generally cut to about 500 μm thickness, down to final device thickness of about 10 μm. Also further fabrication steps (polishing the wafers, mounting them on suitable substrates and polishing them to the final device thickness) was labour intensive.

The epitaxial techniques offer, in comparison with bulk growth techniques, the possibility of growing large area epilayers and sophisticated device structures with good lateral homogeneity and abrupt and complex composition and doping profiles, which can be configured to improve the performance of the photodetectors. The growth is performed at low temperatures (see figure 22), which makes it possible to reduce the native defects density. The properties of HgCdTe grown by the variety of techniques discussed here are summarized in table 3.

Among the various epitaxial techniques, the liquid phase epitaxy (LPE) is the most mature method. LPE is a single crystal growth process in which growth from a cooling solution occurs on a substrate. At present the vapour phase epitaxy (VPE) growth of HgCdTe is typically done by nonequilibrium methods; metalorganic chemical vapour deposition (MOCVD), molecular beam epitaxy (MBE) and their derivatives. The great potential benefit of MBE and MOCVD over the equilibrium methods is the ability to modify the growth conditions dynamically during growth to tailor band gaps, add and remove dopants, prepare surfaces and interfaces, add passivations, perform anneals and even grow on selected areas of a substrate. The growth control is exercised with great precision to obtain basic materials properties comparable to those routinely obtained from equilibrium growth.

Epitaxial growth of HgCdTe layers requires a suitable substrate. CdTe was used initially, since it was available from commercial sources in reasonably large sizes. The main drawback of

Table 3. Comparison of the various methods used to grow HgCdTe (after [48]).

	Bulk						
		Travelling heater method		Liquid phase epitaxy		Vapour phase epitaxy	
	SSR	HCT melt	Te melt	Hg melt	Te melt	MOCVD	MBE
Temperature (°C)	950	950	500	350–550	400–550	275–400	160–200
Pressure (Torr)	150 000	150 000	760–8000	760–11 400	760–8000	300–760	10^{-3}–10^{-4}
Growth rate (μm h^{-1})	250	250	80	30–60	5–60	2–10	1–5
Dimensions w (cm)	0.8–1.2 dia	0.8–1.2 dia	2.5 dia	5	5	7.5 dia	7.5 dia
l (cm)	—	—	—	6	5	4	4
t (cm)	15	15	15	0.0002–0.0030	0.0005–0.012	0.0005–0.001	0.0005–0.001
Dislocations (cm^{-2})	$<10^5$	—	$<10^5$	$<10^5$	$<10^5$–10^7	5×10^5–10^7	$<5 \times 10^4$–10^6
Purity (cm^{-3})	$<5 \times 10^{14}$	$<5 \times 10^{14}$	$<5 \times 10^{14}$	$<5 \times 10^{14}$	$<5 \times 10^{14}$	$<1 \times 10^{15}$	$<1 \times 10^{15}$
n-type doping (cm^{-3})	N/A	N/A	N/A	1×10^{14}–1×10^{18}	1×10^{15}–1×10^{16}	5×10^{14}–5×10^{18}	5×10^{14}–1×10^{19}
p-type doping (cm^{-3})	N/A	N/A	N/A	1×10^{15}–1×10^{18}	1×10^{15}–5×10^{16}	3×10^{15}–5×10^{17}	1×10^{16}–5×10^{18}
X-ray rocking curve (arc sec)	—	—	20–60	<20	<20	50–90	20–30
Compositional uniformity (Δx)	<0.002	<0.004	<0.005	<0.002	<0.002	±0.01–0.0005	±0.01–0.0006

CdTe is that it has a lattice mismatch with LWIR and MWIR HgCdTe of a few per cent . By the mid-1980s it was demonstrated that the addition of a few per cent of ZnTe to CdTe (typically 4%) could create a lattice-matched substrate. CdTe and closely lattice-matched CdZnTe substrates are typically grown by the modified vertical and horizontal unseeded Bridgman technique. Most commonly the (111) and (100) orientations have been used, although others have been tried such as (110) and (123). Twinning, which occurs in (111) layers, can be prevented by a suitable disorientation of the substrate. The limited size, purity problems, Te precipitates, dislocation density (routinely in the low $10^4\,\mathrm{cm}^{-2}$ range), non-uniformity of lattice match and the high price (\$60–\$500 per $1\,\mathrm{cm}^2$, polished) are the remaining problems required to be solved. It is believed that these substrates will continue to be important for a long time, particularly for the highest performance devices.

The LPE growth of thin layers on CdTe substrates began in the early-to-mid 1970s. Initially, Te solutions with dissolved Cd (Cd has a high solubility in Te) and saturated with Hg vapour were used to efficiently grow HgCdTe in the temperature range of 420–600°C. This allowed small-volume melts to be used with slider techniques which did not appreciably deplete during the growth run. Experiments with Hg-solvent LPE began in the late 1970s. Because of the limited solubility of Cd in Hg, the volume of the Hg melts had to be much larger than the Te melts (typically about 20 kg) in order to minimize melt depletion during layer growth in temperature ranges of 380–500°C. This precluded the slider growth approach and Hg-melt epitaxy has been developed using large dipping vessels. Four layers ($30\,\mathrm{cm}^2$ each) with a total area of $120\,\mathrm{cm}^2$ can be grown in a single run [64].

In the early 1990s, bulk growth was replaced by LPE and is now very mature for production of first- and second-generation detectors. The LPE technique has been successfully used in fabrication of DLHJ p-on-n photodiodes (see figure 19), where a p^+ cap layer is formed using arsenic as the dopant. However, LPE technology is limited for a variety of advanced HgCdTe structures required for third generation detectors. LPE typically melts off a thin layer of the underlying material, and each time, an additional layer is grown as a result of relatively high growth temperature. Additionally, the gradient in x-value in the base layer of p^+-on-n junction can generate a barrier transport in certain cases due to interdiffusion. These limitations have provided an opportunity for vapour phase epitaxy: MBE and MOCVD.

The era of MBE and MOCVD began in the early 1980s by adopting both methods well established in the III–V semiconductor materials. Through the next decade a variety of metalorganic compounds were developed along with a number of reaction-chamber designs [65]. In the case of MBE, a specially designed Hg-source ovens was successfully designed to overcome the low sticking coefficient of Hg at the growth temperature [66]. The growth temperature is less than 200°C for MBE but around 350°C for MOCVD, making it more difficult to control the p-type doping in the MOCVD due to the formation of Hg vacancies at higher growth temperatures. At present, MBE is the dominant vapour phase method for HgCdTe. It offers low temperature growth under an ultrahigh vacuum environment, *in situ* n-type and p-type doping, control of composition and doping and interfacial profiles. MBE is now the preferred method for growing complex layer structures for multi-colour detectors and for avalanche photodiodes. Although the quality of MBE material is not yet on par with LPE, it has made tremendous progress in the past decade. The keys of this success have been the doping ability and the reduction of etch pit densities (EPDs) to below $10^5\,\mathrm{cm}^{-2}$.

Near lattice matched CdZnTe substrates have severe drawbacks such as the lack of large areas, high production costs and more importantly, the difference in the thermal expansion coefficient (TEC) in CdZnTe substrates and silicon readout integrated circuits, as well as interest in large-area based IR FPAs (1024×1024 and larger) have resulted in limitations of CdZnTe substrate application. Currently, readily producible CdZnTe substrates are limited

to areas of approximately 30 cm^2. At this size, the wafers are unable to accommodate more than two 1024 × 1024 FPAs. Not even a single die can be printed for very large FPA formats (2048 × 2048 and larger) on substrates of this size.

The use of Si substrates is very attractive in IR FPA technology, not only because it is less expensive and available in large area wafers but also because the coupling of the Si substrates with Si readout circuitry in an FPA structure allows fabrication of very large arrays exhibiting long-term thermal cycle reliability. Despite the large lattice mismatch (≈19%) between CdTe and Si, MBE has been successfully used for the heteroepitaxial growth of CdTe on Si. Using optimized growth conditions for Si(211)B substrates, CdTe(211)B layers with EPD of 10^6 cm^{-2} range could be obtained [67]. While this value of EPD has little effect on MWIR HgCdTe/Si detectors, it must be reduced to achieve high-performance LWIR detectors. By comparison, HgCdTe epitaxial layers grown by MBE or LPE on bulk CdZnTe have typical EPD value in the 10^4 to mid-10^5 cm^{-2} range where there is a negligible effect of dislocation density on detector performance [68]. Defect densities for LWIR material grown on silicon substrates continue to limit performance but they have been reduced by an order of magnitude in the past decade. Recently published results report growth of LWIR HgCdTe epilayers on 4-inch Si substrates with dislocation density in the mid-10^6 cm^{-2} range [69, 70]. At 77 K, diode performance with a cutoff wavelength of about 10 μm is comparable to that of LWIR HgCdTe on bulk CdZnTe substrates [70].

Sapphire has also been widely used as a substrate for HgCdTe epitaxy. In this case a CdTe (CdZnTe) film is deposited on the sapphire prior to the growth of HgCdTe. This substrate has excellent physical properties and can be purchased in large wafer sizes. Large lattice mismatch with HgCdTe is accommodated by a CdTe buffer layer. The typical defect density of HgCdTe on sapphire is in the mid-10^5 cm^{-2} range. Sapphire is transparent from the UV to about 6 μm and has been used to backside-illuminate SWIR and MWIR detectors (it is not acceptable for backside-illuminated LWIR arrays because of the opacity beyond 6 μm).

6. HgCdTe photodiodes

Different HgCdTe photodiode architectures have been fabricated that are compatible with backside and frontside illuminated hybrid FPA technology. The most important architectures are included in table 4 which summarizes the applications of HgCdTe photodiode designs by the major FPA manufacturers today.

The p–n junctions have been formed by numerous techniques including Hg in- and out-diffusion, impurity diffusion, ion implantation, electron bombardment, plasma induced type conversion, doping during growth from vapour or liquid phase and other methods.

The low binding energies and ionic bond nature of HgCdTe give rise to two important effects, which are influential in most junction forming processes. The first is the role of Hg, which is liberated readily by processes such as ion implantation and ion beam milling. This creates a much deeper junction than would be expected from the implantation range. A second effect is the role of dislocations, which may play a part in annihilating vacancies. The role of Hg interstitials, dislocations and ion bombardment in the junction forming process is complex and not well understood in detail. Despite the complex physics involved, manufacturers have received good phenomenological control of the junction depth and dopant profiles with a variety of processes. Of late epitaxial techniques with doping during growth are most often used for preparing p-on-n junctions. MBE and MOCVD have been successfully accomplished with As doping during growth.

Figure 23 shows the schematic band profiles of the most commonly used unbiased homo- (n^+-on-p) and heterojunction (p-on-n) photodiodes. To avoid contribution of the tunnelling

current, the doping concentration in the base region below $10^{16}\,\mathrm{cm}^{-3}$ is required. In both photodiodes, the lightly doped narrow gap absorbing region 'base' of the photodiode: p(n)-type carrier concentration of about $5 \times 10^{15}\,\mathrm{cm}^{-3}$ ($5 \times 10^{14}\,\mathrm{cm}^{-3}$) determines the dark current and photocurrent. The internal electric fields at interfaces are 'blocking' for minority carriers and the influence of surface recombination is eliminated. Also, suitable passivation prevents the influence of surface recombination. Indium is most frequently used as a well-controlled dopant for n-type doping due to its high solubility and moderately high diffusion. Elements of the VB group are acceptors substituting Te sites. They are very useful for fabrication of stable junctions due to very low diffusivity. Arsenic proved to be the most successful p-type dopant to date. The main advantages are stability in lattice, low activation energy and the possibility of controlling concentration over the 10^{15}–$10^{18}\,\mathrm{cm}^{-3}$ range. Intensive efforts are currently underway to reduce the high temperature (≈400°C) required to activate As as an acceptor.

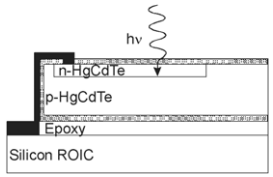
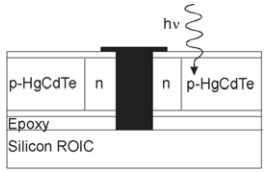
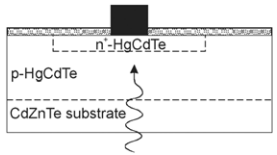
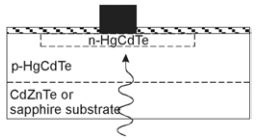
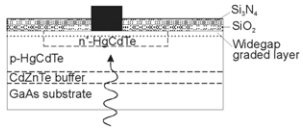
The key technology needed to make photodiodes possible was surface passivation. Based on silicon's success passivation efforts were initially focused mainly on oxides (see figure 11). Anodic oxide was adequate for photoconductors due to a fixed positive charge. Applied to photodiodes, anodic oxide shorted out the devices by inverting the p-type surface. Silicon oxide was employed for photodiode passivation in the early 1980s based upon low-temperature deposition using a photochemical reaction. It appears, however, that the excellent surface properties (with low state densities and excellent photodiode properties) could not be maintained when the device was heated in vacuum for extended periods of time, a procedure required for good vacuum packaging integrity [48]. Also surface charge buildup was created when operated in a space-radiation environment. Recent efforts are concentrated mostly on passivation with CdTe and CdZnTe. Much of the pioneering work in this area was initially done in France at Societe Anonyme de Telecommunication (SAT) in the mid-1970s [80, 81]. CdTe passivation is stable during vacuum packaging bake cycles and shows little effect from the radiation found in space applications. Diodes do not show a variation in R_0A product with diode size, indicating that surface perimeter effects can be neglected.

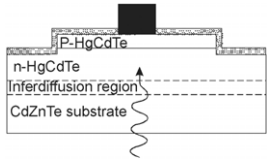
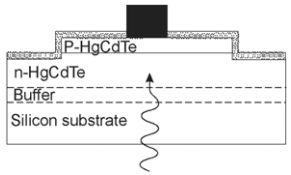
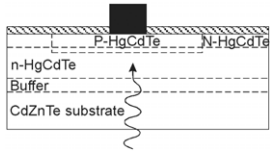
In ideal photodiodes the diffusion current is dominant, therefore their leakage current is very low and insensitive to the detector bias. Leakage current is the primary contribution of unwanted noise. Figure 24 shows typical I–V characteristics of an HgCdTe photodiode at temperatures between 40 and 90 K for a 12 μm cutoff detector at 40 K. The leakage current is less than $10^{-5}\,\mathrm{A\,cm}^{-2}$ at 77 K. The bias-independent leakage current makes it easier to achieve better FPA uniformity, as well as to reduce the detector bias-control requirements during changes in photocurrent.

The quality of HgCdTe photodiodes has improved steadily over the past twenty years as materials and device processing science evolved and has progressed to the point where there is usually no clear indication of g–r current. The plots of R_0A data versus temperature generally follow a diffusion current dependence at higher temperatures, and the transition into a comparatively temperature-independent tunnelling-like regime at lower temperatures. An example of such behaviour is shown in figure 25(a), where the R_0A product versus temperature is presented under 0° FOV for a top-quality 10 μm cutoff wavelength boron-implanted n^+-n^--p HgCdTe photodiode. Figure 25(b) shows the R_0A product versus temperature for a variety p-on-n $Hg_{1-x}Cd_xTe$ photodiodes made from a range of alloy compositions. The longer wavelength devices are typically more difficult to produce than medium or short wavelength diodes.

Figure 26 is an accumulation of R_0A-data taken on P-n-n HgCdTe DLHJ photodiodes (capital letter means wider gap region) with different cutoff wavelength [83]. The solid line is calculated from equation (36) for 1D diffusion current for an n-type absorber layer

Table 4. HgCdTe photodiode architectures used for hybrid FPAs.

Configuration	Architecture	Junction formation	Company	References
n-on-p VIP		Ion implantation forms n-on-p diode in p-type HgCdTe, grown by Te-solution LPE on CdZnTe and epoxied to silicon ROIC wafer; over the edge contact	DRS Infrared Technologies (formerly Texas Instruments)	[71]
n–p loophole		Ion beam milling forms n-type islands in p-type Hg-vacancy-doped layer grown by Te-solution LPE on CdZnTe, and epoxied onto silicon ROIC wafer; cylindrical lateral collection diodes	GEC-Marconi Infrared (GMIRL)	[72]
n^+-on-p planar		Ion implant into acceptor-doped p-type LPE film grown by Te-solution slider	Sofradir (Societe Francaise de Detecteurs Infrarouge)	[73]
n^+-n^--p planar homojunctions		Boron implant into Hg-vacancy p-type, grown by Hg-solution tipper on 3 inch diameter sapphire with MOCVD CdTe buffer; ZnS passivation	Rockwell/Boeing	[74]
n^+-n^--p planar homojunctions		MBE grown of n-type layer on GaAs substrate with CdZnTe buffer layer. Boron implant into converted p-type layer. SiO_2/Si_3N_4 passivation.	Institute of Semiconductor Physics, Novosibirsk	[75]

P-on-n mesa		1. Two-layer LPE on CdZnTe: Base: Te-solution slider, indium-doped Cap: Hg-solution dipper, arsenic-doped 2. MOCVD *in situ* on CdZnTe Iodine-doped base, arsenic-doped cap	IR Imaging Systems, Sanders—A Lockheed Martin Company (LMIRIS)	[76] [77]
P-on-n mesa		1. Two-layer LPE on CdZnTe or Si: Base: Hg-solution dipper, indium-doped Cap: Hg-solution dipper, arsenic-doped 2. MBE *in situ* on CdZnTe or Si Indium-doped base, arsenic-doped cap	Raytheon Infrared Center of Excellence (RIRCoE, formerly SBRC) and Hughes Research Laboratories (HRLs)	[78] [79]
P-on-n planar buried heterostructure		Arsenic implant into indium-doped N-n or N-n-N film grown by MBE on CdZnTe	Rockwell/Boeing	[74]

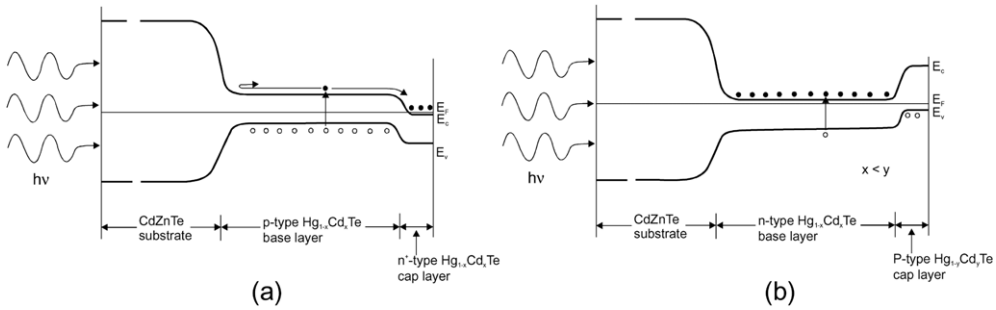


Figure 23. Schematic band diagrams of n$^+$-on-p homojunction (*a*) and p-on-n heterojunction (*b*) photodiodes.

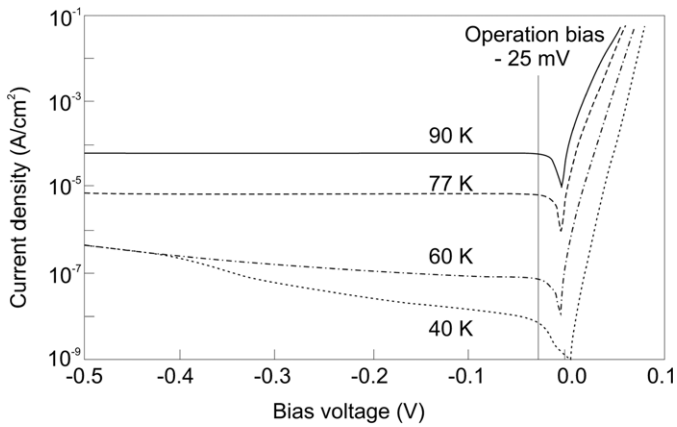


Figure 24. *I–V* characteristics at various temperatures for a 12 μm cutoff HgCdTe photodiode (after [82]).

with only Auger 1 and radiative recombination. The dashed line is the theoretical $(R_0A)_{\rm opt}$, also calculated for the 1D model. Typical values for the n-side donor concentration ($N_{\rm d} = 1 \times 10^{15}\,{\rm cm}^{-3}$), the narrow band gap active layer thickness and the detector implant area ($50 \times 5\,\mu{\rm m}^2$) were used in theoretical calculations. The experimental values in the figure were determined from the measured R_0 from *I–V* characteristics and the optical area $A_{\rm opt}$ determined from measured flood and focused spot quantum efficiency. The data show that the $(R_0A)_{\rm opt}$ values calculated from measured values of R_0 are equal to, or within a factor of two, below the theoretical curve for the LWIR and VLWIR detectors. The MWIR detectors are above one order of magnitude below the theoretical curve.

Several major advantages in the development of VLWIR HgCdTe technology have been recently presented by Reine *et al* [60, 84]. Continued improvements in defect reduction have enabled R_0A products to follow the diffusion current limit down to temperatures of 40–45 K for cutoff wavelengths up to 19 μm (see figure 27).

The Auger 1 and Auger 7 recombination mechanisms are relatively insignificant in the wide gap HgCdTe alloys needed for SWIR and MWIR applications; thus, the only fundamental recombination mechanism to be considered is radiative recombination. This is illustrated in figure 28 for three groups of p-on-n devices fabricated using LPE, MBE and MOCVD. For photodiodes at temperature 180 K, the R_0A data generally fall about a factor of ten

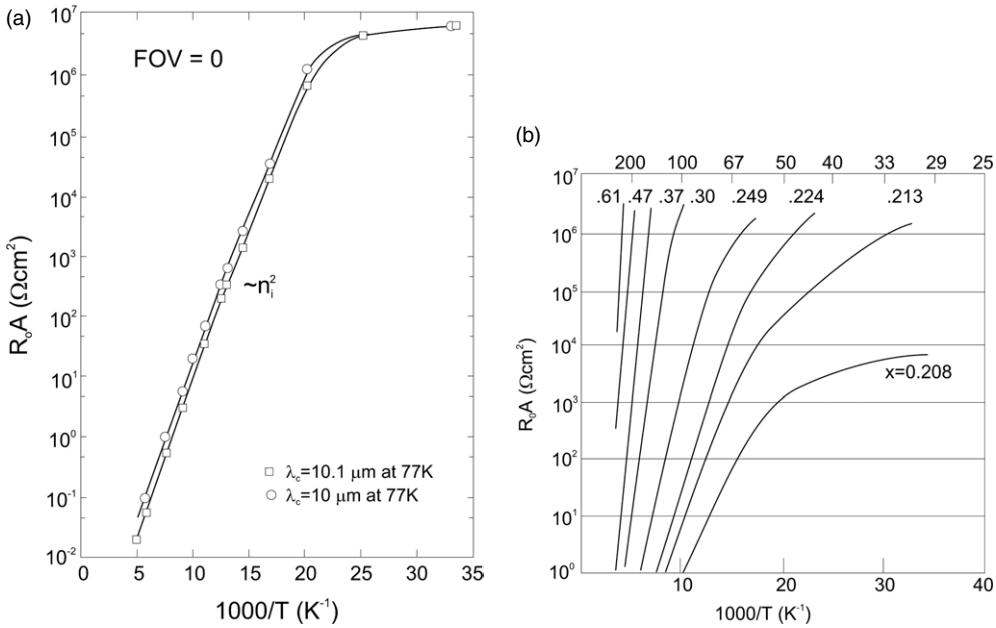


Figure 25. R_0A product versus temperature for HgCdTe photodiodes: (*a*) 10 μm cutoff wavelength boron-implanted n^+-n^--p photodiode (after [20]); (*b*) a variety p-on-n $Hg_{1-x}Cd_xTe$ photodiodes made from a range of alloy compositions, as illustrated (after [48]).

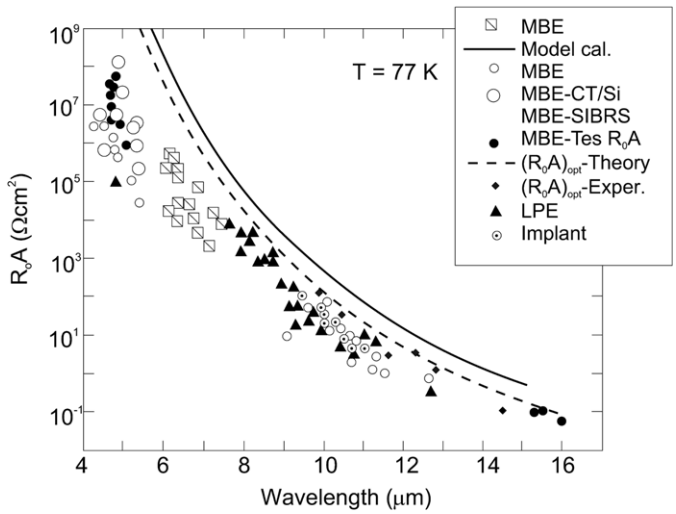


Figure 26. R_0A versus cutoff wavelength for p-on-n HgCdTe photodiodes at 77 K (after [83]).

below the theoretical curves, indicating that a lifetime mechanism other than the traditional radiative recombination is lowering the lifetime. According to DeWames *et al* [87], a shallow SR recombination centre, possibly process-induced, is responsible for the reduced lifetime. It appears, that for photodiodes operated at room temperature with cutoff wavelengths less than 3.5 μm, the R_0A product falls short of the limits calculated with traditional radiative

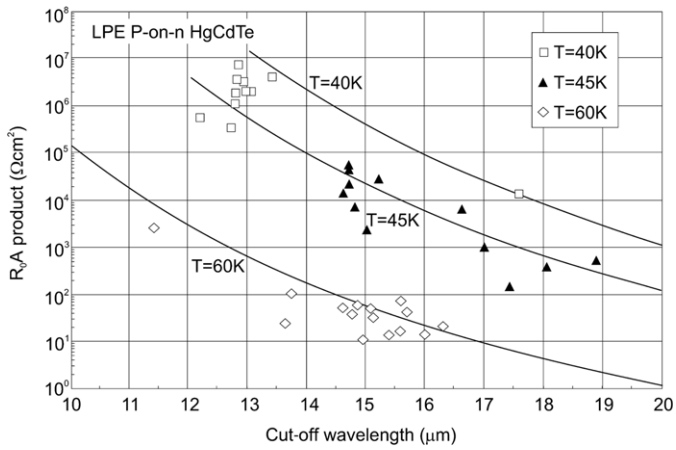


Figure 27. Median R_0A data for 20-element arrays of LPE p-on-n HgCdTe photodiodes at 60, 45 and 40 K, compared with calculated R_0A due to n-side diffusion current (——) (after [60]).

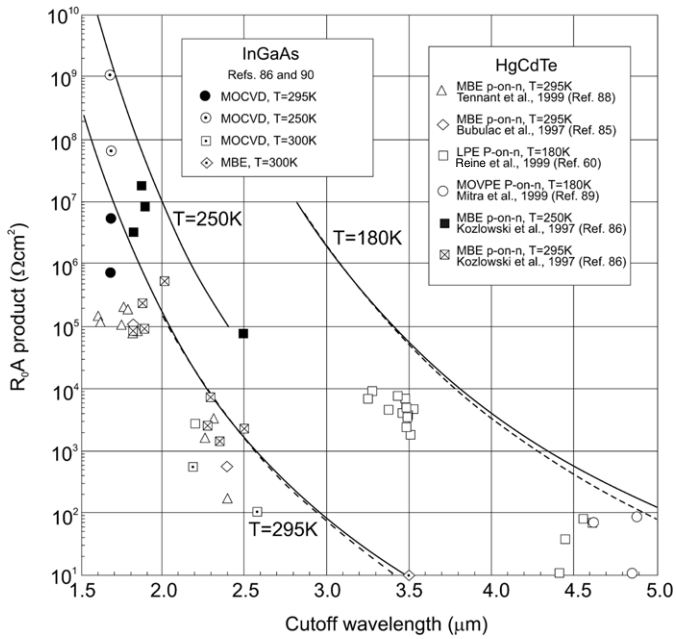


Figure 28. R_0A data for SWIR HgCdTe and InGaAs photodiodes, all with n-type absorber layers, compared with calculated R_0A due to n-side diffusion current with only radiative (——) and radiative and Auger 1 (- - - -).

recombination equations. However, the highest quality SW photodiodes fabricated with HgCdTe (and also with InGaAs alloys) have performance levels in agreement with the radiative limit.

Many additional excess mechanisms affect the dark current of HgCdTe photodiodes [3]. They arise from non-fundamental sources located in the base and cap layer, the depletion region and the surface. As the operating temperature is lowered, the thermal dark current mechanisms

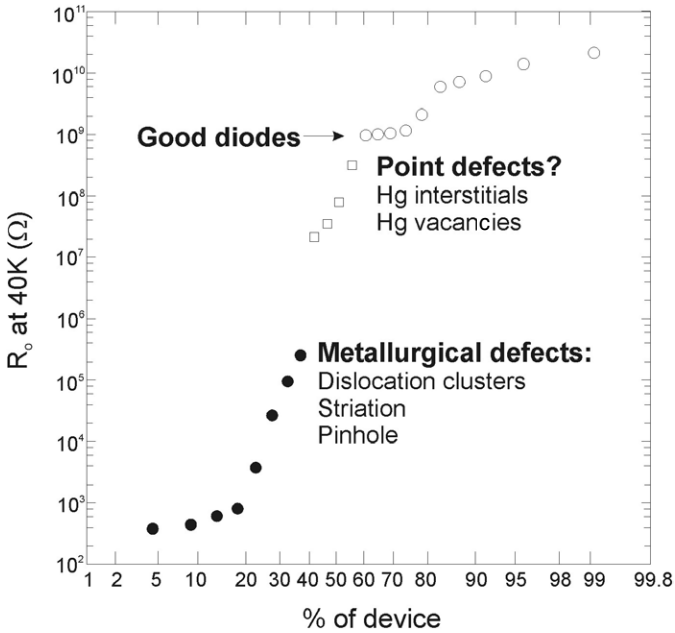


Figure 29. Detailed analysis separates the cumulative distribution function of R_0A values of LWIR p-on-n HgCdTe photodiodes (fabricated by LPE) into three regions: good diodes, diodes affected by point defects and diodes affected by metallurgical defects (after [91]).

become weaker and allow other mechanisms to prevail. In practice, non-fundamental sources dominate the dark current of the present HgCdTe photodiodes, with the exception of specific cases of near room temperature devices and the highest quality 77 K LWIR/VLWIR and 200 K MWIR devices. The main leakage mechanisms of HgCdTe photodiodes are generation in the depletion region, interband tunnelling, trap-assisted tunnelling and impact ionization. Some of these are caused by structural defects in the p–n junction. These mechanisms receive much attention now, particularly because they determine ultimately the array uniformity, yield and cost for some applications, especially those with lower operating temperatures.

In $Hg_{1-x}Cd_xTe$ photodiodes with $x \approx 0.22$, the diffusion current in the zero-bias and the low-bias region is usually the dominant current down to 60 K [3]. At medium values of reverse bias, the dark current is mostly due to trap-assisted tunnelling. Trap-assisted tunnelling dominates the dark current also at zero bias below 50 K. At high values of reverse bias, bulk band-to-band tunnelling dominates. At low temperatures, such as 40 K, significant spreads in the R_0A product distributions are typically observed due to the onset of tunnelling currents associated with localized defects. Moreover, HgCdTe photodiodes often have an additional surface-related component of the dark current [3], particularly at low temperatures.

Chen *et al* [91] carried out a detailed analysis of the wide distribution of the R_0 values of HgCdTe photodiodes operating at 40 K. Figure 29 shows the cumulative distribution function, R_0, obtained in devices with a cutoff wavelength between 9.4 and 10.5 μm. It is clear that while some devices exhibit a fair operability with R_0 values spanning only two orders of magnitude, other devices show poor operability with R_0 values spanning more than 5–6 orders of magnitude. Lower performance, with R_0 values below $7 \times 10^6\ \Omega$ at 40 K, is usually due to gross metallurgical defects, such as dislocation clusters and loops, pin holes, striations,

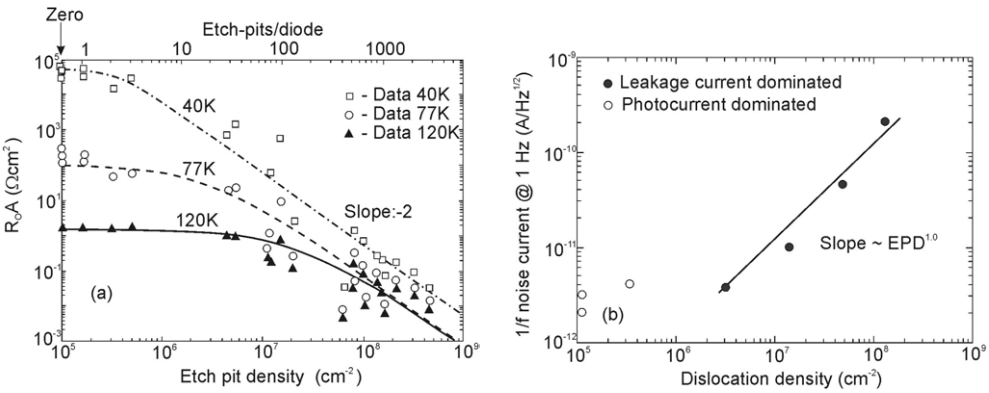


Figure 30. Influence of the dislocation density on the parameters of HgCdTe photodiodes: (*a*) R_0A product versus EPD, showing the fit of model to the data, for a 9.5 μm array (at 78 K), measured at 120, 77 and 40 K at zero FOV; (*b*) $1/f$ noise current at 1 Hz versus dislocation density measured at 78 K for 10.3 μm HgCdTe photodiode array ($f/2$ FOV) (after [93]).

Te inclusions and heavy terracing. However, diodes with R_0 values between $7 \times 10^6\ \Omega$ and $1 \times 10^9\ \Omega$ at 40 K contained no visible defects (Hg interstitials and vacancies).

Dislocations are known to increase the dark current and the $1/f$ noise current. At 77 K the requirement on dislocation density for LWIR material is $<2 \times 10^5\ \mathrm{cm}^{-2}$. MWIR, on the other hand, can tolerate higher densities of dislocations at 77 K, but evidence is mounting that this is no longer true at higher operating temperatures [92]. The reverse bias characteristics of HgCdTe diodes depend strongly on the density of dislocations intercepting the junction. Johnson *et al* [93] showed that in the presence of high dislocations densities the R_0A product decreases at the square of the dislocation density; the onset of the square dependence occurs at progressively lower dislocation densities as the temperature decreases, which is shown in figure 30(*a*). At 77 K, R_0A begins to decrease at an EPD of approximately $10^6\ \mathrm{cm}^{-2}$, while at 40 K R_0A is immediately affected by the presence of one or more dislocations in the diode. The scatter in the R_0A data at large EPD may be associated with the presence of an increased number of pairs of 'interacting' dislocations in some of those diodes; these pairs are more effective in reducing the R_0A than individual dislocations. To describe the dependence of the R_0A product with dislocation density a phenomenological model was developed, which was based on the conductances of individual and interacting dislocations which shunt the p–n junctions. As figure 30(*a*) shows, this model was found to give a reasonable fit to the experimental data.

In general, the $1/f$ noise appears to be associated with the presence of potential barriers at the contacts, interior or surface of the semiconductor. Different models have been proposed to explain experimental data including the modulation of the surface generation current by the fluctuations of the surface potential and the influence of trap-assisted tunnelling across a pinched-off depletion region [3]. Johnson *et al* [93] presented the effect of dislocations on the $1/f$ noise. Figure 30(*b*) shows that, at low EPD, the noise current is dominated by the photocurrent, while at higher EPD the noise current varies linearly with EPD. It appears that dislocations are not the direct source of the $1/f$ noise, but rather increase this noise only through their effect on the leakage current. The $1/f$ noise current varies as $I^{0.76}$ (where I is the total diode current); similar to the fit of the data taken on undamaged diodes.

The average value of the R_0A product at 77 K for a 10 μm cutoff HgCdTe photodiode at 77 K is around $300\ \Omega\,\mathrm{cm}^2$ and drops to $30\ \Omega\,\mathrm{cm}^2$ at 12 μm [3, 60, 83]. At 40 K, the R_0A product varies between 10^5 and $10^8\ \Omega\,\mathrm{cm}^2$ with 90% above $10^5\ \Omega\,\mathrm{cm}^2$ at 11.2 μm.

Table 5. Representative staring HgCdTe hybrid FPAs offered by some major manufacturers.

Manufacturer/ Web site	Size	Pixel size (μm)	Spectral range (μm)	Oper. temp. (K)	$D^*(\lambda_p)$ (cm Hz$^{-1/2}$ W^{-1})/ NETD (mK)
Raytheon/	128 × 128	40 × 40	9–11	80	—
www.raytheon.com	256 × 256	30 × 30	8.5–11	77–100	—
Rockwell/Boeing/	256 × 256	40 × 40	>15	77	—
www.boeing.com	640 × 480	27 × 27	>10	77	—
www.rsc.rockwell.com	225 × 256	40 × 40	1–4.6	120	>10
	640 × 480	27 × 27	1–4.6	120	>25
	2048 × 2048	18 × 18	1–2.5	95–120	$>1 \times 10^{14}$
	4096 × 4096	18 × 18	1–2.5	95–120	$>1 \times 10^{14}$
BAE Systems/	256 × 256	30 × 30	8–10	80	—
www.infrared-detectors.com	384 × 288	20 × 20	3–5	<120	18
	640 × 512	24 × 24	3–5	<110	12
Sofradir/	128 × 128	50 × 50	7.7–10	80	$1.1 \times 10^{11}/10$
www.infrared.sofradir.com	128 × 128	50 × 50	3.7–4.8	90	$4.3 \times 10^{11}/7$
	320 × 256	30 × 30	7.7–9.5	77	⩽20
	320 × 256	30 × 30	3.7–4.8	⩽120	⩽10
	640 × 512	15 × 15	3.7–4.8	⩽110	⩽17
	1000 × 256	30 × 30	0.8–2.5	⩽200	—
DRS Infrared Technologies/	256 × 256	40 × 40	LWIR	77	1.2×10^{10} (D^* 500)
www.drs.com	480 × 640	25 × 25	LWIR	77	3.5×10^{9} (D^* 500)
	256 × 256	40 × 40	MWIR	77	5.5×10^{10} (D^* 500)
	480 × 640	25 × 25	MWIR	77	4.0×10^{10} (D^* 500)
AIM-AEG Infrarot-	128 × 128	40 × 40	8–10	80	<15
Module GmbH/	128 × 128	40 × 40	3–5	80	<15
www.aim-ir.com	256 × 256	40 × 40	8–10	80	<20
	256 × 256	40 × 40	3–5	80	<10
	384 × 288	24 × 24	8–10	80	<40
	384 × 288	24 × 24	3–5	80	<25
	640 × 512	24 × 24	3–5	80	<18

Up to the present, photovoltaic HgCdTe FPAs have been mainly based on p-type material. Linear (240, 288, 480 and 960 elements), 2D scanning arrays with TDI (with common formats of 256 × 4, 288 × 4, 480 × 6) and 2D starring formats from 64 × 64 up to 2048 × 2048 have been made with a wide range of spectral response [58]. In the last ten years however, considerable progress in the fabrication of p-on-n photodiodes has been achieved. The performance of p-on-n photodiodes is even better than the n-on-p ones. Efforts are also underway to develop avalanche photodiode capabilities in the 1.6 μm and at longer wavelength region. Pixel sizes ranging from 18 μm square to over 1 mm have been demonstrated.

The best results have been obtained using hybrid architecture. However, the very large hybrid structures suffer from mechanically fatigue problems that result from different thermal expansion between the detector and readout structures. Therefore, considerable effort is being extended to develop techniques for growing HgCdTe on silicon substrates.

Table 5 contains a description of representative starring HgCdTe FPAs that are commercially available as standard products and/or catalogue items from the major manufacturers.

Rockwell has developed the world's largest HgCdTe SWIR FPA for astronomy and low background applications (see figure 31) [94, 95]. The format of the device is a hybrid

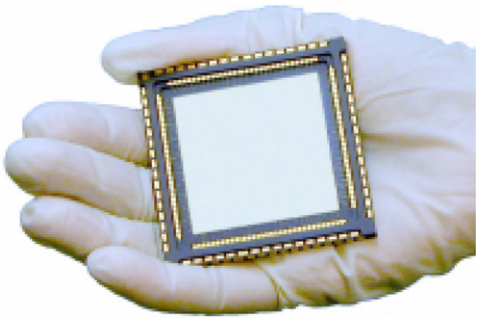


Figure 31. A 2048 × 2048 SWIR (1–3 μm) HgCdTe Hawaii array with 18 μm pixels (after [58]).

2048 × 2048 with a unit cell size of 18 × 18 μm^2 and with active size of 37 mm. Sets of four arrays will be 'tiled' into 2 × 2 mosaic configurations giving 4096 × 4096 pixels [96]. One of the 4096 × 4096 mosaic will be installed in a camera instrument on Mauna Kea, Hawaii, and the other mosaic will be installed in an IR imager on Gemini South in Chile. The development of large format, high sensitivity mosaic IR sensors for ground-based astronomy is the goal of many observatories around the world (large arrays dramatically multiply the data output of a telescope system). This is somewhat surprising given the comparative budgets of the defence market and the astronomical community.

7. Third generation detectors

In the last decade of the twentieth century (see figure 11) a third generation of HgCdTe detectors emerged from the tremendous impetus in the detector developments. The definition of the third generation IR systems is not particularly well established. In the common understanding, third generation IR systems provide enhanced capabilities like a larger number of pixels, higher frame rates, better thermal resolution as well as multi-colour functionality and other on-chip functions. According to Reago *et al* [97] the third generation is defined in order to maintain the current advantage enjoyed by the US and allied armed forces. This class of devices includes both cooled and uncooled FPAs [97, 98]:

- high performance, high resolution cooled imagers having multi-colour bands,
- medium- to high-performance uncooled imagers,
- very low cost, expendable uncooled imagers.

Many challenges have forced the IR community to develop third generation IR imagers. Some of them are considered here:

- noise equivalent difference temperature (NEDT),
- pixel and chip size issues,
- uniformity and
- identification and detection ranges.

7.1. Noise equivalent difference temperature

For FPAs the relevant figure of merit is the NEDT, the temperature change of a scene required to produce a signal equal to the rms noise. It can be shown that [99]

$$\mathrm{NEDT} = (\tau C \eta_{\mathrm{BLIP}} \sqrt{N_{\mathrm{w}}})^{-1}, \tag{41}$$

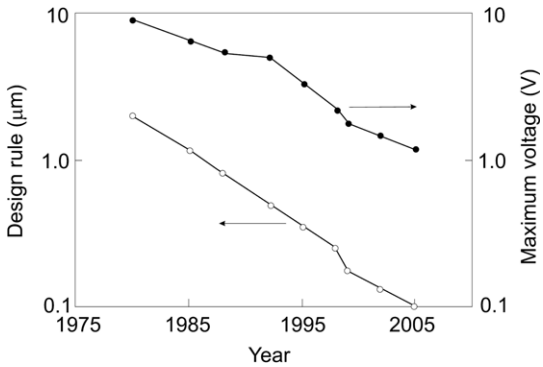


Figure 32. Trends for design rule minimum dimensions and maximum bias voltage of silicon foundry requirements (after [98]).

where τ is the optics transmission spectrum and C is the thermal contrast. N_w is the number of photogenerated carriers integrated for one integration time, t_{int}, and Φ_B is the photon flux density incident on detector area A_d

$$N_w = \eta A_d t_{int} \Phi_B. \tag{42}$$

The percentage of BLIP, η_{BLIP}, is simply the ratio of photon noise to composite FPA noise

$$\eta_{BLIP} = \left(\frac{N_{photon}^2}{N_{photon}^2 + N_{FPA}^2} \right)^{1/2}. \tag{43}$$

It results from the above formulae that the charge handling capacity of the readout, the integration time linked to the frame time and dark current of the sensitive material becomes the major issues of IR FPAs. The NEDT is inversely proportional to the square root of the integrated charge and therefore the greater the charge, the higher the performance.

Second generation imagers provide NEDT of about 20–30 mK with $f/2$ optics. A goal of third-generation imagers is to achieve sensitivity improvement corresponding to NEDT of about 1 mK. It results from equation (42) that in a 300 K scene in LWIR region with thermal contrast of 0.04, the required charge storage capacity is above 10^9 electrons. This high charge-storage density within the small pixel dimensions probably cannot be obtained with standard CMOS capacitors [98]. Although the reduced oxide thickness of submicrometre design rules gives large capacitance per unit area, the reduced bias voltage, as illustrated in figure 32, largely cancels any improvement in charge storage density. Ferroelectric capacitors may provide much greater charge storage densities than the oxide-on-silicon capacitors now used; however, such technology is not yet incorporated into standard CMOS foundries. Norton *et al* [98] have suggested using stacked hybrid structures at least as an interim solution to incorporate the desired charge storage density in detector–readout–capacitor structures.

7.2. Pixel and chip sizes

Pixel and chip sizes are important issues in the association with multi-colour imager formats. Small pixels reduce cost by increasing the number of readout and detector dice potentially available from processed wafers. Small pixels also allow smaller, lighter optics to be used.

The fundamental limit of the pixel size is determined by diffraction. The size of diffraction-limited optical spot or Airy disk is given by

$$d = 2.44\lambda f, \tag{44}$$

where d is the diameter of the spot, λ is the wavelength and f is the f-number of the focusing lens. For typical $f/2.0$ optics at 5 μm wavelength, the spot size is 25 μm. Because the system users prefer some degree of oversampling, the pixel size may be reduced for MWIR applications of the order of 12 μm. Given the track record of human nature, Norton [100] has predicted that MWIR pixel size will eventually be reduced to about 10 μm at some point, just to achieve the smaller pixel record. SWIR pixel size will shrink to correspondingly smaller dimensions, for applications seeking maximum spatial resolution. LWIR pixels should not shrink much below 20 μm. However, it is anticipated that LWIR pixels will be made as small as MWIR pixels, since a single readout design may be used with both MWIR and LWIR FPAs.

Readout wafers are processed in standard commercial foundries and can be constrained in size by the die-size limits of the photolithography step and repeat printers [98]. This limit is currently of the order of $22 \times 22\,\mathrm{mm}^2$ for submicron lithography. Then, the array itself can only occupy $18 \times 18\,\mathrm{mm}^2$ assuming one needs about 2 mm on each side for periphery circuitry (bias supplies, shift registers, column amplifiers and output drivers). In these conditions, a 1024×1024 array would need to have pixels no larger than 18 μm.

The 1024×1024 Hawaii readout, for example, can be made using a GCA wafer stepper; its area encompasses $3.8\,\mathrm{cm}^2$ and a total of >3.4 M transistors. Some newer steppers capable of supporting 0.25 μm processes have slightly larger fields. For the largest HgCdTe FPA with a unit cell size of $18 \times 18\,\mu\mathrm{m}^2$, the 2048×2048 CMOS readout has been fabricated by abutting four optical fields to create each circuit. Each quadrant is accurately placed to $<0.05\,\mu$m to fabricate a $4 \times 4\,\mathrm{cm}^2$ readout [94].

It will be an extremely difficult challenge to deploy a two- or three-colour detector structure into a small pixel such as $18 \times 18\,\mu\mathrm{m}^2$. Current two-colour simultaneous mode pixels with two indium bumps per pixel have not been built with pixels less than 30 μm.

7.3. Uniformity

As discussed by Levine [101], when the detectivity is approaching a value above $10^{10}\,\mathrm{cm\,Hz^{1/2}\,W^{-1}}$, the FPA performance is uniformity limited prior to correction and thus essentially independent of the detectivity. An improvement in non-uniformity from 0.1% to 0.01% after correction could lower the NEDT from 63 to 6.3 mK.

The non-uniformity value is usually calculated using the standard deviation over mean, counting the number of operable pixels in an array. For a system operating in the LWIR band, the scene contrast is about $2\%\,\mathrm{K}^{-1}$ of change in the scene temperature. Thus, to obtain a pixel-to-pixel variation in apparent temperature to less than, e.g. 20 mK, the non-uniformity in response must be less than 0.04%. This is nearly impossible to obtain in the uncorrected response of the FPA, so a two-point correction is typically used.

FPA uniformity influences an IR system complexity. The uniformity is important for accurate temperature measurements, background subtraction and threshold testing. Non-uniformities require elaboration of compensation algorithms to correct the image and by consuming a number of analog-to-digital bits they also reduce the system dynamic range.

Tactical IR FPAs usually require operating in the LWIR window with a small number of applications in the 3–5 μm window. Ranges from the sensor to the target are typically short allowing the use of imaging sensors with large FPAs where precise radiometry is not critical. Imaging arrays can usually tolerate some percentage of dead or degraded pixels without jeopardizing mission performance. Tactical backgrounds in the IR windows are relatively high with about 10^{16} photons $\mathrm{cm^{-2}\,s^{-1}}$ reaching the detector.

Figure 1 plots the energy band gap, $E_g(x, T)$, for $Hg_{1-x}Cd_xTe$ versus alloy composition x for temperature of 77 K. Also plotted is the cutoff wavelength $\lambda_c(x, T)$, defined as that

Table 6. Cutoff wavelength for x variations of 0.2% and the corresponding cutoff wavelength shift for $Hg_{1-x}Cd_xTe$.

Composition, x	Cutoff wavelength λ_c (μm)	Temperature T (K)	Uncertainty $\Delta\lambda_c$ (μm)
0.395	3	77	0.023
0.295	5	77	0.064
0.210	10	77	0.260
0.196	14	77	0.510

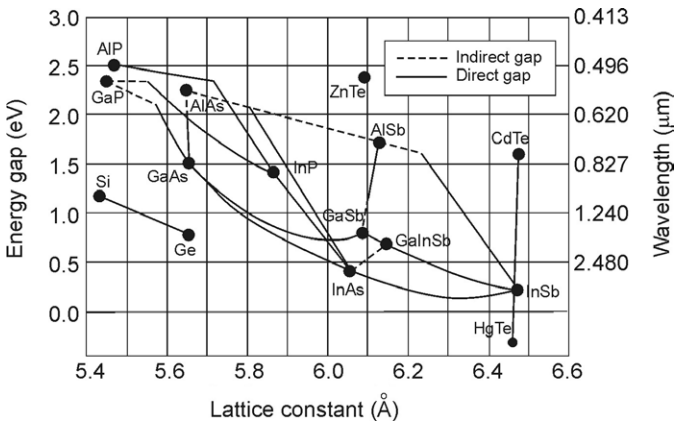


Figure 33. Composition and wavelength diagram of Sb-based III–V material systems.

wavelength at which the response has dropped to 50% of its peak value. Table 6 shows uncertainty in cutoff wavelength for x variations of 0.2%. For SWIR ($\approx$3 μm) and MWIR ($\approx$5 μm) materials, the variation in cutoff wavelength is not large. However, the non-uniformity is a serious problem in the case of LWIR HgCdTe detectors. The variation of x across the $Hg_{1-x}Cd_xTe$ wafer causes much larger spectral nonuniformity; e.g. at 77 K, a variation of $\Delta x = 0.2\%$ gives a $\Delta\lambda_c = 0.064\,\mu$m at $\lambda_c = 5\,\mu$m, but $\Delta\lambda_c = 0.51\,\mu$m at 14 μm) which cannot be fully corrected by the two or three point corrections [102]. Therefore, required composition control is much more stringent for LWIR than for MWIR. For applications that require operation in the LWIR band as well as two-colour MWIR/LWIR bands, most probably HgCdTe will not be the optimal solution.

The alternative candidates for third generation IR detectors are Sb-based III–V material systems. These materials are mechanically robust and have fairly weak dependence of band gap on composition (see figure 33).

7.4. Identification and detection ranges

A thermal imaging system is used to first detect an object, and then identify it. Typically, identification ranges are between two and three times shorter than detection ranges [103]. To increase ranges, increasing resolution and sensitivity of the IR systems (and hence the detectors) are required. Third generation cooled imagers are being developed to extend the range of target detection and identification and to ensure that army forces maintain a technological advantage in night operations over any opposing force.

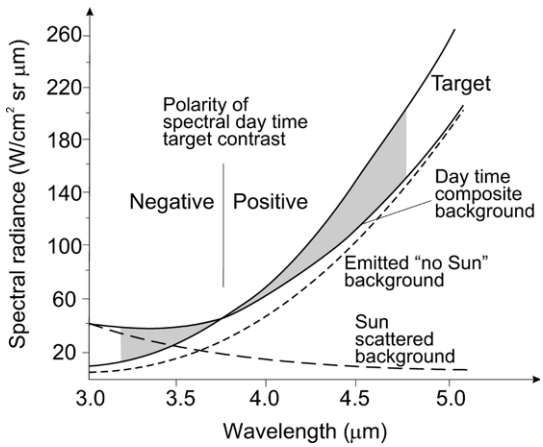


Figure 34. Target and background contrast reversal in the MWIR spectral range (after [103]).

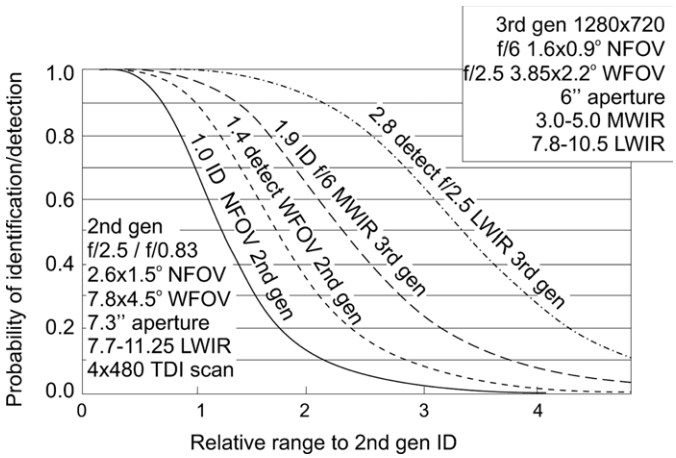


Figure 35. Comparison of the detection and identification range between current second generation TDI scanned LWIR imagers and the LWIR and MWIR bands of third generation imager in a 1280×720 format with 20 μm pixels (after [104]).

Identification ranges can be further increased by using multispectral detection to correlate the imagers in different wavelengths. For example, it appears that in an MWIR spectral range the IR image is washed out to the point that the target and the background cannot be distinguished from each other (see figure 34). Detectors that cover the entire spectral range will get washed out because the background contrast changes from positive to negative. Alternatively, using two band detectors (up to 3.8 μm and from 3.8 up to 5 μm) and summing up the inverse of the second band and the output of the first band, will yield a contrast enhancement unlike that if an integrated response of the entire spectral range was used.

Figure 35 compares the relative detection and identification ranges modelled for third generation imagers using NVESD's (Fort Belvoir, VA 22060) NVTherm program. As a range criterion, the standard 70% probability of detection or identification is assumed. Note that the identification range in the MWIR range is almost 70% of the LWIR detection range. For

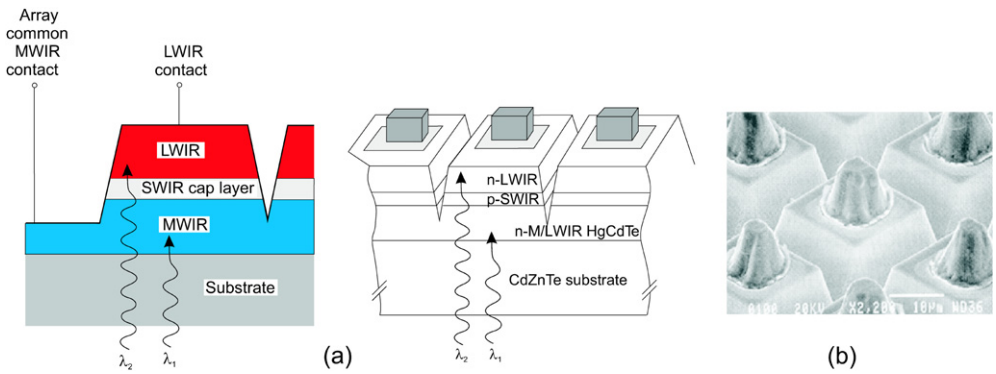


Figure 36. Two-colour n-P-N detector structure for sequential operating mode: (*a*) cross section schematic and (*b*) a SEM image.

detection, LWIR provides a superior range. In the detection mode, the optical system provides a wide FOV (WFOV-$f/2.5$) since third generation systems will operate as an on-the-move wide area step-scanner with automated target recognition (second generation systems rely on manual target searching) [104]. MWIR offers higher spatial resolution sensing and advantages for long-range identification when used with telephoto optics (NFOV-$f/6$).

7.5. Two-colour HgCdTe detectors

Considerable progress in the development of multi-colour HgCdTe FPAs has been recently demonstrated by research groups at HRL, Raytheon, Lockheed Martin (BAE Systems), DRS IR Technology, AIM, Leti, Rockwell and NVESD employing mainly MBE (although LPE and MOCVD are also used) for the growth of a variety of devices [105–111].

The unit cell of integrated two-colour FPAs consists of two co-located detectors, each sensitive to a different spectral band. In the back-illuminated dual-band detectors, the photodiode with longer cutoff wavelength is grown epitaxially on top of the photodiode with the short cutoff wavelength. The shorter cutoff photodiode acts as a long-wavelength-pass filter for the longer cutoff photodiode.

Both sequential mode and simultaneous mode detectors are fabricated from the multilayer materials. The simplest two-colour HgCdTe detector, and the first to be demonstrated, is the bias-selectable n-P-N triple-layer heterojunction (TLHJ), back-to-back photodiode shown in figure 36 (capital letter means wider band gap structure). The n-type base absorbing regions are deliberately doped with indium at a level of about (1–3) $\times$ 10^{15} cm^{-3}. A critical step in device formation is connected with the *in situ* p-type As-doped layer with good structural and electrical properties to prevent internal gain from generating spectral crosstalk. The band gap engineering effort consists of increasing the CdTe mole fraction and the effective thickness of the p-type layer to suppress out-of-band carriers from being collected at the terminal. The sequential-mode detector has a single indium bump per unit cell that permits sequential bias-selectivity of the spectral bands associated with operating back-to-back photodiodes. When the polarity of the bias voltage applied to the bump contact is positive, the top (LW) photodiode is reverse biased and the bottom (SW) photodiode is forward biased. The SW photocurrent is shunted by the low impedance of the forward-biased SW photodiode, and the only photocurrent to emerge in the external circuit is the LW photocurrent. When the bias voltage polarity is reversed, the situation reverses; only SW photocurrent is available. Switching times within the

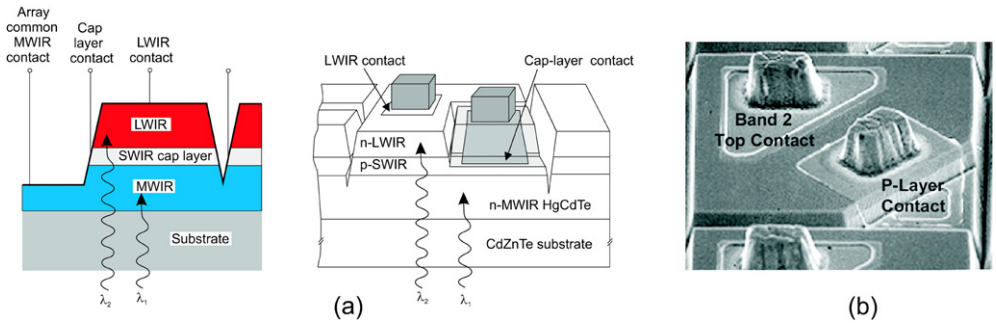


Figure 37. Two-colour n-P-N detector structure for simultaneous operating mode: (*a*) cross section schematic and (*b*) a SEM image.

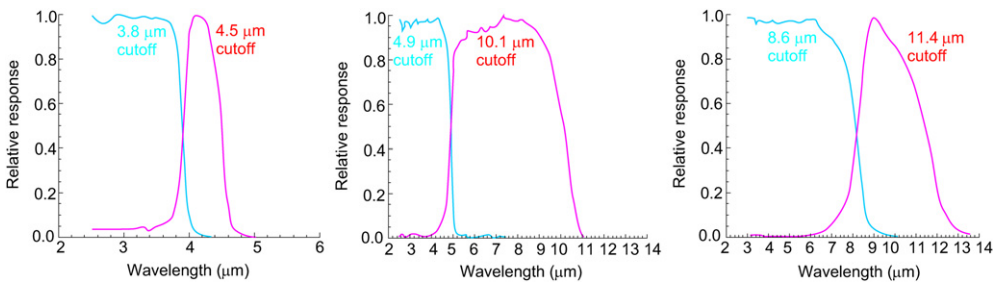


Figure 38. Spectral response curves for two-colour HgCdTe detectors in various dual-band combinations of MWIR and LWIR spectral bands (after [100]).

detector can be relatively short, in the order of microseconds, so detection of slowly changing targets or images can be done by switching rapidly between the MW and LW modes.

One bump contact per unit cell, as for single-colour hybrid FPAs, is the big advantage of the bias-selectable detector. It is compatible with existing silicon readout chips. The problems with the bias selectable device are the following: its construction does not allow independent selection of the optimum bias voltage for each photodiode and there can be substantial MW crosstalk in the LW detector.

Many applications require true simultaneity of detection in the two spectral bands. Simultaneous dual-band detector architectures require an additional electrical contact to an underlying layer in the multijunction structure of both the SW and LW photodiodes. An implementation of the simultaneous mode using a second indium bump in the unit cell is shown in figure 37. The most important distinction is the requirement of a second readout circuit in each unit cell.

Integrated two-colour HgCdTe technology has been developed for more than a decade with steady progression having been made with a wide variety of pixel sizes (30–61 μm), array formats (64×64 up to 320×240) and spectral-band sensitivity (MWIR/MWIR, MWIR/LWIR and LWIR/LWIR). Figure 38 shows examples of spectral response from different two-colour devices. Note that there is minimal crosstalk between the bands, since the short wavelength band absorbs nearly 100% of the shorter wavelengths. Test structures indicate that the separate photodiodes in a two-colour detector perform exactly as single-colour detectors in terms of achievable R_0A product variation with wavelength at a given temperature (see table 7).

Figure 39 presents the NEDT for both bands as a function of temperature for simultaneous mode two-colour 128×128 MWIR/MWIR FPAs. Fill factors as high as 80% were achieved

Table 7. Typical measured performance parameters for single- and dual-colour HgCdTe MWIR and LWIR detector configuration for 256 × 256 30 μm unit-cell FPAs (after [110]).

256 × 256 30 μm unit-cell performance parameters	DLHJ single colour		TLHJ sequential dual colour					
	MWIR	LWIR	MWIR/MWIR		MWIR/LWIR		LWIR/LWIR	
Spectral band	MWIR	LWIR	Band 1	Band 2	Band 1	Band 2	Band 1	Band 2
78 K cutoff (μm)	5	10	4	5	5	10	8	10
Operating temperature (K)	78	78	120	120	70	70	70	70
Cross talk (%)	—	—	<5	<10	<5	<10	<5	<10
Quantum efficiency (%)	>70	>70	>70	>65	>70	>50	>70	>50
R_0A, zero FOV ($\Omega\,\mathrm{cm}^2$)	$>1 \times 10^7$	>500	6×10^5	2×10^5	1×10^6	2×10^2	5×10^4	5×10^2
R_rA[a], zero FOV ($\Omega\,\mathrm{cm}^2$)	—	—	—	—	—	—	—	—
Interconnect operability (%)	>99.9	>99.9	>99.9	>99.9	>99.9	>99.9	>99.9	>99.9
Response operability (%)	>99	>98	>99	>97	>99	>97	>98	>95

[a] Resistance area product at nonzero bias.

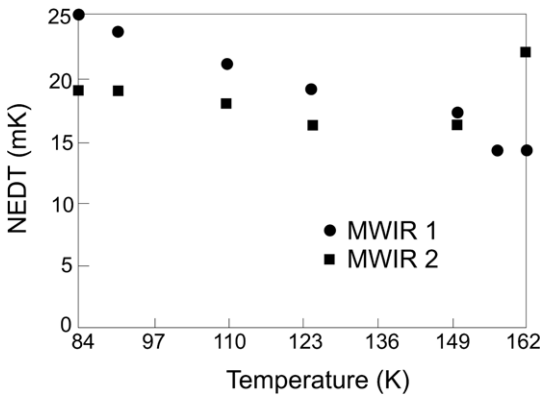


Figure 39. Noise-equivalent difference temperature for two-colour camera having 50 mm, $f/2.3$ lens, as a function of operating temperature (after [82]).

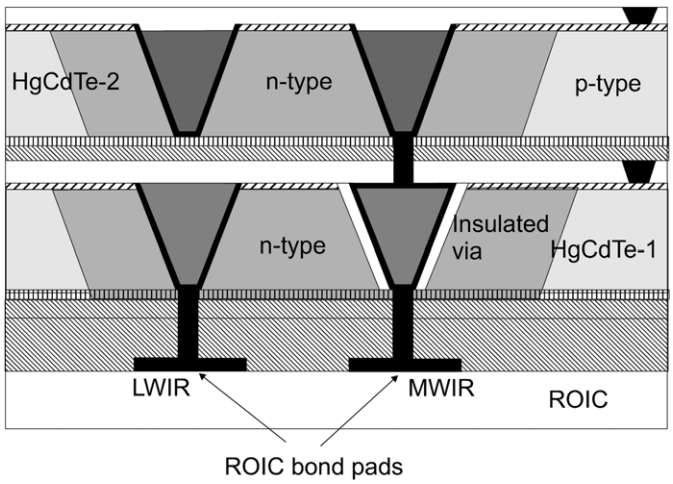


Figure 40. Schematic representation of DRS HDVIP^{TM} process used to produce dual-band HgCdTe FPA (after [108]).

using a single mesa structure to accommodate the two indium bump contacts required for each unit cell with 50 μm size. Band 1 (2.5–3.9 μm) had operability of 99.9%, with 23 inoperable pixels. Band 2 (3.9–4.6 μm) had operability of 98.9%, with 193 inoperable pixels. Quantum efficiencies of 70% were observed in each band without using an anti-reflection coating. The R_0A values for the diodes ranged from 8.25×10^5 to 1.1×10^6 $\Omega\,\mathrm{cm}^2$ at $f/2$ FOV. The camera used for these measurements had a 50 mm, $f/2.3$ lens. Imagery was acquired at temperatures as high as 180 K with no visible degradation in image quality.

The HgCdTe high density vertically integrated (HDVIP^{TM}) or loophole concept (see figure 17), developed at DRS and BAE Southampton, represents an alternative approach to IR FPA architecture. It differs from the more entrenched FPA architectures in both its method of diode formation and the manner of its hybridization to the silicon ROIC. This architecture has been extended into two-colours at DRS by gluing two mono-colour layers together into a composite, and forming an insulated via through the lower layer in order to read out the signal of upper colour layer as illustrated in figure 40. The ROIC used for the dual-band FPA was

originally designed for a single-colour 640 × 480 array with 25 μm (square) pixels. The even numbered rows of the ROIC have no detectors attached to them so the chip is operated in a mode that only outputs the odd rows. Odd numbered columns connect to LWIR detectors and the MWIR detectors are on the even columns. This approach has been utilized to fabricate both MW/LW and MW/MW 240 × 320 FPAs on a 50 μm pitch. Higher densities are being investigated with dedicated two-colour ROIC designs, enabling pitches of <30 μm for two-colour FPAs.

8. Alternative material systems

The difficulties in growing HgCdTe material significantly due to solidus–liquidus separation (leading to marked segregation between CdTe and HgTe) and the high vapour pressure of Hg, encouraged the development of alternative technologies over the past 40 years. One of these was PbSnTe, a IV–VI compound material system which was vigorously pursued in parallel with HgCdTe in the late 1960s and early 1970s [112, 113]. InAs/$Ga_{1-x}In_xSb$ SLSs have been also proposed for IR detector applications in the 8–14 μm region [35]. Among different types of quantum well IR photodetectors (QWIPs) technology of the GaAs/AlGaAs multiple quantum well detectors is the most mature. The QWIP technology is relatively new and has been developed very quickly in the last decade [34, 101, 114–116] with LWIR imaging performance comparable to state-of-the-art of HgCdTe. The technologies competitive to HgCdTe ternary alloy are compared in the following section.

8.1. Lead salt ternary alloys

Lead salt ternary alloys (PbSnTe and PbSnSe) seemed easier to prepare and appeared more stable. Development of IV–VI alloy photodiodes was discontinued because the chalcogenides suffered two significant drawbacks. The first was a high dielectric constant that resulted in high diode capacitance and therefore limited frequency response (for PbSnTe, the observed values of the static dielectric constant have been widely distributed from 400 to 5800, and at the same temperature these values have been scattered in the range up to one order of magnitude [16, 117]). For scanning systems under development at that time, this was a serious limitation. However, for starring imaging systems under development today, using 2D arrays, this would not be as significant an issue.

The second drawback to IV–VI compounds is their very high thermal coefficients of expansion (TEC) [118]. This limited their applicability in hybrid configurations with silicon multiplexers (differences in TEC between the readout and detector structure can lead to failure of the indium bonds after repeated thermal cycling from room temperature to the cryogenic temperature of operation). Today, with the ability to grow these materials on alternative substrates such as silicon, this too would not be a fundamental limitation [119, 120].

In practice, the response speed of a photodiode is determined by the effects of junction capacitance, dynamic resistance and series resistance together with external circuit impedance. If we assume that the only significant capacitance is the junction space–charge region capacitance C, and the only resistance is the external load resistance R_L, then the upper-frequency limits f_c is

$$f_c = \frac{1}{2\pi R_L C}. \tag{45}$$

Figure 41 shows plots of the cutoff frequency versus applied reverse-bias voltage for one-side abrupt n–p$^+$ $Pb_{0.78}Sn_{0.22}Te$ and $Hg_{0.797}Cd_{0.203}Te$ photodiodes at 77 K ($\lambda_c \approx 12\,\mu$m).

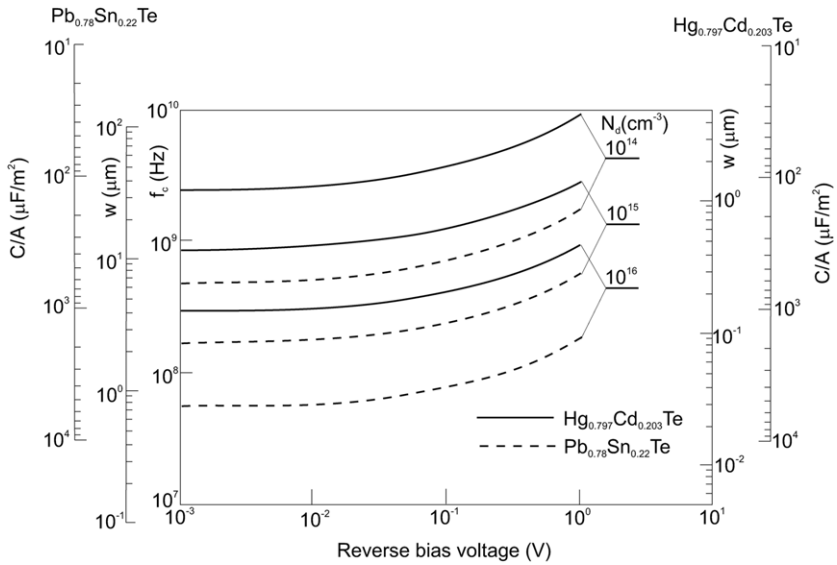


Figure 41. Cutoff frequency for one-side abrupt n–p$^+$ $Pb_{0.78}Sn_{0.22}Te$ and $Hg_{0.797}Cd_{0.203}Te$ photodiodes ($\lambda_c \approx 12\,\mu$m) at 77 K with an active area of 10^{-4} cm^2. The additional scales correspond to the depletion width and the junction capacitance per unit area (after [44]).

For reverse-bias voltage above 1 V and doping concentrations in the space charge region above 10^{15} cm^{-3}, an avalanche breakdown can be observed. The cutoff frequency has been calculated for a load resistance of 50 Ω and for a junction area of 10^{-4} cm^2. Various values of donor concentrations are marked in figure 41. In this figure the space charge region width and the capacitance per unit, C/A, are also shown. We can see that cutoff frequencies of 2 GHz can be realized with HgCdTe photodiodes when n-side doping concentration is not larger than 10^{14} cm^{-3} at reverse-bias voltages. The cutoff frequency for PbSnTe photodiodes is almost an order of magnitude smaller. It should be noted that the cutoff frequency is decreased by any junction series resistance and stray capacitance.

Figure 42 shows dependence of the TEC of PbTe, InSb, HgTe and Si on temperature. At room temperature, the TCE HgTe and CdTe is about 5×10^{-6} K^{-1}, while that of PbSnTe is in the range of 20×10^{-6} K^{-1}. This results in a much greater TCE mismatch with silicon (TCE about 3×10^{-6} K^{-1}). It might be noted that both Ge and GaAs have TCE values close to HgCdTe, giving detectors on those materials no significant advantage in this respect.

The doping concentrations in the base regions of HgCdTe and PbSnTe photodiodes are different. For both types of photodiodes, the tunnelling current (and R_0A product) is critically dependent on doping concentration. Figure 9 shows the dependence of the R_0A product components on the dopant concentrations for one-sided abrupt HgCdTe, PbSnTe and PbSnSe photodiodes at 77 K ($E_g = 0.1$ eV). To produce high R_0A products for HgCdTe and lead salt photodiodes, the doping concentration of 10^{16} and 10^{17} cm^{-3} (or less) are required, respectively.

8.2. InSb and InGaAs

In the middle and late 1950s it was discovered that InSb had the smallest energy gap of any semiconductor known at that time and its applications as a MWIR detector became obvious. The energy gap of InSb is less well matched to the 3–5 μm band at higher operating

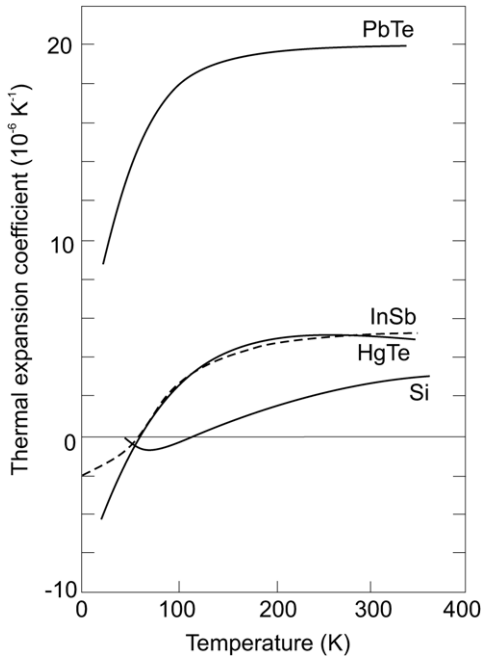


Figure 42. Linear TEC of PbTe, InSb, HgTe and Si versus temperature (after [121]).

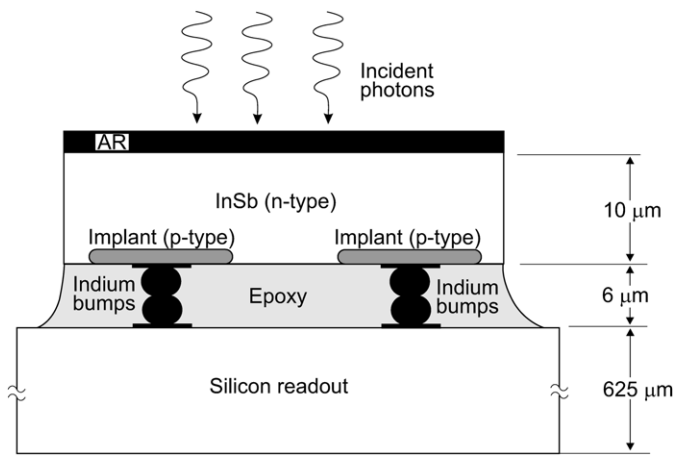


Figure 43. Architecture of an InSb sensor chip assembly (after [122]).

temperatures, and better performance can be obtained from $Hg_{1-x}Cd_xTe$. InAs is a compound similar to InSb, but has a larger energy gap, so that the threshold wavelength is 3–4 μm.

In InSb photodiode fabrication, epitaxy is not used; instead, the standard manufacturing technique begins with bulk n-type single crystal wafers with donor concentration of about 10^{15} cm^{-3}. Relatively large bulk grown crystals with 2 and 3 inch diameters are available in the market. An array hybrid size up to 2048×2048 is possible because the InSb detector material is thinned to less than 10 μm (after surface passivation and hybridization to a readout chip) which allows it to accommodate the InSb/silicon thermal mismatch [122]. As is shown in figure 43,

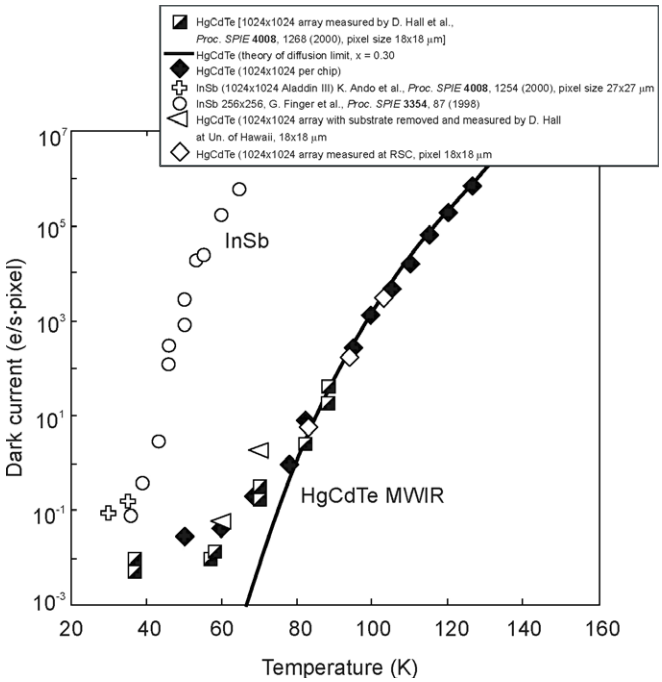


Figure 44. The comparison of dependence of dark current on temperature between MBE-grown MWIR FPAs and highest reported value for InSb arrays. The HgCdTe 1024 × 1024 arrays with 18 × 18 μm pixels. The HgCdTe cutoff is 5.3 μm and no AR coating, quantum efficiency is 73% at 78 K (after [124]).

the backside illuminated InSb p-on-n detector is a planar structure with an ion-implanted junction. After hybridization, epoxy is wicked between the detector and the Si ROIC and the detector is thinned to 10 μm or less by diamond-point turning. One important advantage of a thinned InSb detector is that no substrate is needed; these detectors also respond to the visible portion of the spectrum. Recently, the growth of InSb and related alloys by MBE together with the doping of the substrate to induce transparency has also been demonstrated [123]. In the last case the thinning of the detector material is not required.

Figure 44 compares the dependence of dark current on temperature between HgCdTe and InSb photodiodes. This comparison suggests that MWIR HgCdTe photodiodes have significantly higher performance in the 30–120 K temperature range. The InSb devices are dominated by generation–recombination currents in the 60–120 K temperature range because of a defect centre in the energy gap, whereas MWIR HgCdTe detectors do not exhibit g–r currents in this temperature range and are limited by diffusion currents. In addition, wavelength tunability has made HgCdTe the preferred material.

$In_{0.53}Ga_{0.47}As$ alloy ($E_g = 0.73$ eV, $\lambda_c = 1.7$ μm) lattice matched to the InP substrate has already been shown to be a suitable detector material for the near-IR (1.0–1.7 μm) spectral range. Having lower dark current and noise than indirect-band gap germanium, the competing near-IR material, the material addresses both entrenched applications including lightwave communication systems, low-light-level night vision and new applications such as remote sensing, eye-safe range finding and process control [3]. Due to the similar band structure of InGaAs and HgCdTe ternary alloys, the ultimate fundamental performance of both types of photodiodes are similar in the wavelength range $1.5 < \lambda < 3.7$ μm [90]. Figure 28 compares

Table 8. Essential properties of LWIR HgCdTe photodiodes and QWIPs at 77 K.

Parameter	HgCdTe	QWIP (n-type)
IR absorption	Normal incidence	$E_{\text{optical}} \perp$plane of well required. Normal incidence: no absorption
Quantum efficiency	$\geqslant$70%	$\leqslant$10%
Spectral sensitivity	Wide-band	Narrow-band (FWHM $\approx$ 1–2 μm)
Optical gain	1	0.2 (30–50 wells)
Thermal generation lifetime	$\approx$1 μs	$\approx$10 ps
R_0A product ($\lambda_c = 10\,\mu$m)	300 $\Omega\,\text{cm}^2$	$10^4\,\Omega\,\text{cm}^2$
Detectivity ($\lambda_c = 10\,\mu$m, FOV = 0)	$2 \times 10^{12}\,\text{cm}\,\text{Hz}^{1/2}\,\text{W}^{-1}$	$2 \times 10^{10}\,\text{cm}\,\text{Hz}^{1/2}\,\text{W}^{-1}$

the ultimate performance of n-type base InGaAs with attainable experimental data. InGaAs photodiodes have shown high device performance close to the theoretical limits for material whose composition is nearly matched to that of InP ($\cong$1.7 μm cutoff wavelength) and InAs ($\cong$ 3.6 μm cutoff wavelength). However, the performance of InGaAs photodiodes decreases rapidly at intermediate wavelengths due to mismatch-induced defects.

Standard $In_{0.53}Ga_{0.47}As$ photodiodes have detector-limited room temperature detectivity of $\sim 10^{13}\,\text{cm}\,\text{Hz}^{1/2}\,\text{W}^{-1}$. With increasing cutoff wavelength, detectivity decreases.

The largest and finest pitched imager in $In_{0.53}Ga_{0.47}As$ material system has been demonstrated recently [125]. The 640 $\times$ 512 FPAs with 25 μm pixels is sensitive to the 0.9–1.7 μm and features a room temperature detectivity greater than $5 \times 10^{12}\,\text{cm}\,\text{Hz}^{1/2}\,\text{W}^{-1}$ with greater than 98% of the pixels operable.

8.3. GaAs/AlGaAs QWIPs

An alternative hybrid detector for the LWIR region (8–14 μm) is the QWIPs. This high impedance detector is built by alternating thin layers (SLs) of GaAs and AlGaAs. Despite large research and development efforts, large photovoltaic HgCdTe FPAs remain expensive, primarily because of the low yield of operable arrays. The low yield is due to the sensitivity of LWIR HgCdTe devices to defects and surface leakage, which is a consequence of the basic material properties. With respect to HgCdTe detectors, GaAs/AlGaAs quantum well devices have a number of potential advantages, including the use of standard manufacturing techniques based on mature GaAs growth and processing technologies, highly uniform and well-controlled MBE growth on greater than 6 inch GaAs wafers, high yield and thus low cost, more thermal stability and extrinsic radiation hardness.

LWIR QWIP cannot compete with the HgCdTe photodiode as the single device, especially at higher temperature operations (>70 K) due to fundamental limitations associated with intersubband transitions. QWIP detectors have relatively low quantum efficiencies, typically less than 10%. The spectral response band is also narrow for this detector, with a full-width, half-maximum of about 15%. All the QWIP data with cutoff wavelength about 9 μm is clustered between 10^{10} and $10^{11}\,\text{cm}\,\text{Hz}^{1/2}\,\text{W}^{-1}$ at about 77 K operating temperature. However, the advantage of HgCdTe is less distinct in a temperature range below 50 K due to the problems involved in an HgCdTe material (p-type doping, SR recombination, trap-assisted tunnelling, surface and interface instabilities). A more detailed comparison of both technologies has been given by Rogalski [34] and Tidrow *et al* [102, 126]. Table 8 compares the essential properties of both types of devices at 77 K.

Even though QWIP is a photoconductor, several of its properties such as high impedance, fast response time long integration time and low power consumption comply well with

the requirements of large FPAs fabrication. The main drawbacks of LWIR QWIP FPA technology are the performance limitations for low integration time applications and low operating temperature. Their main advantages are linked to performance uniformity and to availability of large size arrays. The large industrial infrastructure in III–V materials/device growth processing and packaging brought about by the utility of GaAs-based devices in the telecommunications industry gives QWIPs a potential advantage in producibility and cost. The only known use of HgCdTe, to date, is for IR detectors. The main drawback of LWIR HgCdTe FPA technology is the unavailability of large size arrays necessary for TV formats and larger ones.

State-of-the-art QWIP and HgCdTe FPAs provide similar performance figures of merit, because they are predominantly limited by the readout circuits. It can be shown that the NEDT value for charge-limited QWIP detectors is even better than for HgCdTe photodiodes by factor $(2g)^{1/2}$ since a reasonable value of g is 0.4 [34]. A low photoconductive gain actually increases the signal-to-noise ratio and a QWIP FPA can have a better NEDT than an HgCdTe FPA with similar storage capacity. This deduction was experimentally confirmed by Schneider *et al* [127]. Using a photovoltaic 'low-noise' QWIP structure, in which g is only 0.05, the group achieved an NEDT of 7.4 mK with 20 ms integration time and 5.2 mK with 40 ms.

The very short integration time of LWIR HgCdTe devices of typically below 300 μs is very useful to freeze a scene with rapidly moving objects. QWIP devices achieve, due to excellent homogeneity and low photoelectrical gain, an even better NEDT; however, the integration time must be 10 to 100 times longer for that, and typically it is 5–20 ms. The decision of the best technology is therefore driven by the specific needs of a system. The observation of the global market through the past several years has indicated that even HgCdTe photodiodes intrinsically exhibit higher performances than the QWIP detectors and QWIP detectors will be used for large formats (e.g. 1024 × 1024 and larger) with low frame rates and large integration time. Recently, 1 megapixel hybrid corrugated detector array with V-grooves and 18 μm pixel pitch has been demonstrated to cover the 8–14 μm spectral range [128]. This technology can be extended to a 2 × 2K array, but at present the limitation is the readout availability and cost.

Powerful possibilities of QWIP technology are connected with multi-colour detection. A four-band FPA has been demonstrated by stacking different multi-quantum well structures, which are sensitive in 4–6, 8.5–10, 10–12, and 13–15 μm bands [116, 129]. The 640 × 512 format FPA consists of four 640 × 128 pixel areas which are capable of acquiring images in these bands. Four separate detector bands were defined by a deep-trench etch process and the unwonted spectral bands were eliminated by a detector short-circuiting process. The unwonted top detectors were electrically shorted by gold-coated reflective 2D etched gratings as shown in figure 45.

8.4. InAs/GaInSb strained layer superlattices

InAs/$Ga_{1-x}In_xSb$ (InAs/GaInSb) SLSs can be considered an alternative to HgCdTe and GaAs/AlGaAs IR material system as a candidate for third generation IR detectors. The low quantum efficiency of QWIPs is largely due to the fact that the optical transition is forbidden for normal incidence of light. Straylight generated by reflecting gratings is required to achieve reasonable quantum efficiency. On the other hand this straylight degrades the modulation transfer function of QWIPs, since some light intensity is guided by the residual substrate into neighbours. In the case of InAs/GaInSb SLS structures the absorption is strong for the normal incidence of light. Consequently, the SLS structures provide high responsivity, as already reached with HgCdTe, without any need for gratings. Further advantages are a

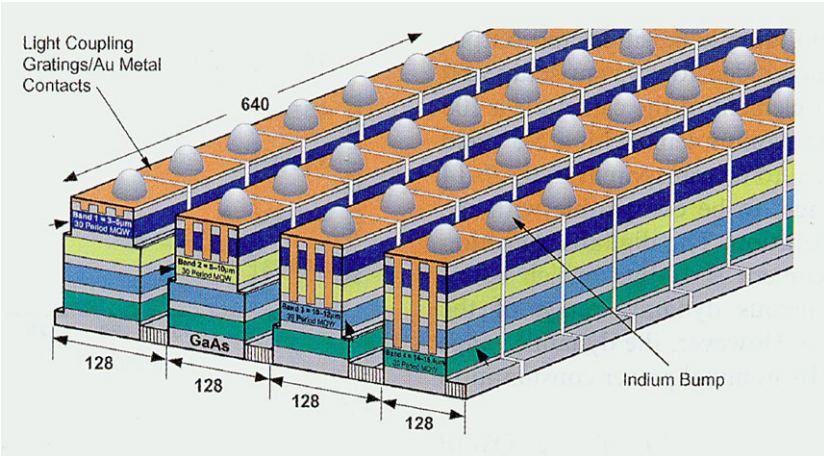


Figure 45. Layer diagram of the four-band QWIP device structure and the deep groove 2D-periodic grating structure. Each pixel represents a 640 × 128 pixel area of the four-band FPA (after [129]).

photovoltaic operation mode, operation at elevated temperatures and well established III–V process technology.

The InAs/GaInSb material system is, however, in a very early stage of development. Problems exist in material growth, processing, substrate preparation and device passivation [130]. Optimization of superlattice (SL) growth is a trade-off between interface roughness, with smoother interfaces at higher temperature, and residual background carrier concentrations, which are minimized on the low end of this range. As grown material has exhibited a moderate residual doping of about 5×10^{15} cm^{-3}.

It has been suggested that InAs/$Ga_{1-x}In_xSb$ SLS can have some advantages over bulk HgCdTe, including lower leakage currents and greater uniformity [39]. Long wavelength response in these SLs arises due to a type II band alignment and internal strain which lowers the conduction band minimum of InAs and raises the heavy-hole band in $Ga_{1-x}In_xSb$ by the deformation potential effect. This reduced band gap is advantageous because longer cutoff wavelengths can be obtained with reduced layer thickness in the strained SL, leading to optical absorption coefficient comparable to that of HgCdTe.

High performance InAs/GaInSb SL photovoltaic was predicted by the theoretical promise of longer intrinsic lifetimes due to the suppression of Auger recombination mechanism [39]. Figure 46 compares the R_0A values of InAs/GaInSb SLS and HgCdTe photodiodes in the long wavelength spectral range. The upper line denotes the theoretical diffusion limited performance corresponding to Auger 7 limitation in p-type HgCdTe material. As can be seen in the figure, the most recent photodiode results for SLS devices rival that of practical HgCdTe devices, indicating that substantial improvement has been achieved in SLS detector development.

In the high temperature range, the performance of LWIR photodiodes is limited by the diffusion process. In temperature range between 80 and 50 K, the photodiodes are depletion-region (generation–recombination) limited. The trap-assisted tunnelling is dominant only at low temperature (<50 K) with almost constant activation trap density (1×10^{12} cm^{-3}). Assuming identical material quality with the same activated trap density, the R_0A values exceeding 100 Ω cm^2 even with a cutoff wavelength of 14 μm can be achieved.

The presented results indicate that fundamental material issues of InAs/GaInSb SLSs fulfill practical realization of high performance FPAs. First 256 × 256 SLS MWIR detectors have

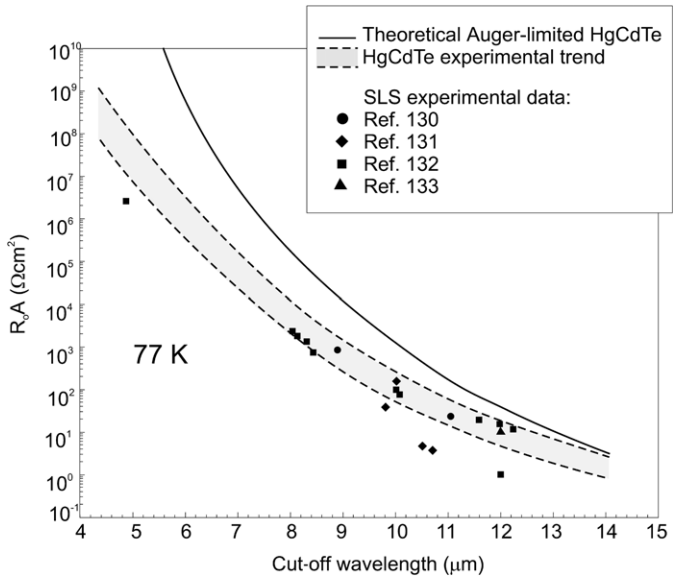


Figure 46. Dependence of the R_0A product of InAs/GaInSb SLS photodiodes on cutoff wavelength compared to theoretical and experimental trendlines for comparable HgCdTe photodiodes at 77 K.

been hybridized [135]. The cutoff wavelength of this detector is 5.3 μm. An excellent NEDT value of approximately 10 mK measured with $f/2$ optics and integration time $\tau_{\text{int}} = 5$ ms has been presented. A very important feature of InAs/GaInSb FPAs is their high uniformity. These very promising results also confirm that the antimonide SLS technology is now competing with MBE HgCdTe dual colour technology.

8.5. Hg-based alternatives to HgCdTe

Among the small gap II–VI semiconductors for IR detectors, only $Hg_{1-x}Zn_xTe$ (HgZnTe) and $Hg_{1-x}Mn_xTe$ (HgMnTe) [136] can be considered as alternatives to HgCdTe. However, both ternary alloy systems have never been systematically explored in the device context. The reasons for this are several. Preliminary investigations of these alloy systems came on the scene when development of HgCdTe detectors was well on its way. Moreover, the HgZnTe alloy is a more serious technological problem material than HgCdTe. In the case of HgMnTe, Mn is not a group II element, so that HgMnTe is not truly a II–VI alloy. This ternary compound was viewed with some suspicion by those not directly familiar with its crystallographic, electrical and optical behaviour. In such a situation, proponents of parallel development of HgZnTe and HgMnTe for IR detector fabrication encountered considerable difficulty in selling the idea to industry and the funding agencies.

9. HgCdTe versus thermal detectors

IR photodetectors operating in MW and LW IR spectral ranges require cryogenic cooling to achieve useful performance. The need for cooling is a major limitation of IR photodetectors which prevents more widespread use of IR technology. Affordable IR systems, such as IR imaging cameras, require cost-effective detectors that operate without cooling or, at least, at temperatures compatible with long-life power and low cost coolers.

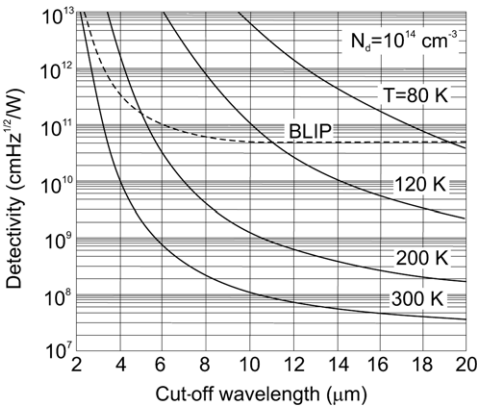


Figure 47. Calculated performance of Auger generation–recombination limited HgCdTe photodetectors as a function of wavelength and temperature of operation. BLIP detectivity has been calculated for 2π FOV, background temperature $T_{\mathrm{BLIP}} = 300\,\mathrm{K}$ and quantum efficiency $\eta = 1$ (after [141]).

The use of thermal detectors operated at room temperature has been the subject of research and development for many decades. Thermal detectors are not useful for high-speed scanning thermal imagers. Only pyroelectric vidicons have found more widespread use. These devices achieved their fundamental limits of performance by about 1970. Generally, the performance of thermal detectors is modest; they suffer from slow response, and they are not useful in applications requiring multispectral detection. However, with large arrays of thermal detectors the best values of NEDT below 0.1 K could be reached because effective noise bandwidths less than 100 Hz can be achieved. This compares with a bandwidth of several hundred kilohertz for conventional cooled thermal imagers with a small photon detector array and scanner. The realization of this fact caused a new revolution in thermal imaging, which is underway now [137–140]. This is due to the development of 2D electronically scanned arrays, in which moderate sensitivity can be compensated for, by a large number of elements. Large scale integration combined with micromachining has been used for manufacturing large 2D arrays of uncooled IR sensors. This enables fabrication of low cost and high-quality thermal imagers. Although developed for military applications, low-cost IR imagers are used in non-military applications such as drivers aid, aircraft aid, industrial process monitoring, community services, firefighting, portable mine detection, night vision, border surveillance, law enforcement, search and rescue, etc. However, thermal detectors cannot replace photon detectors in most applications.

The Auger mechanism is likely to impose fundamental limitations to the LWIR HgCdTe detector performance. Figure 47 shows the calculated detectivity of Auger generation–recombination limited HgCdTe photodetectors as a function of wavelength and temperature of operation. The calculations have been performed for doping level equal $N_{\mathrm{d}} = 10^{14}\,\mathrm{cm}^{-3}$, as the lowest donor doping level, which at present is achievable in a controllable manner in practice. We can see that

- liquid nitrogen cooling potentially makes it possible to achieve BLIP performance in the wide 2–20 μm range,
- 200 K cooling, which is achievable with Peltier coolers, would be sufficient for BLIP operation in the middle and SWIR regions (below 5 μm).

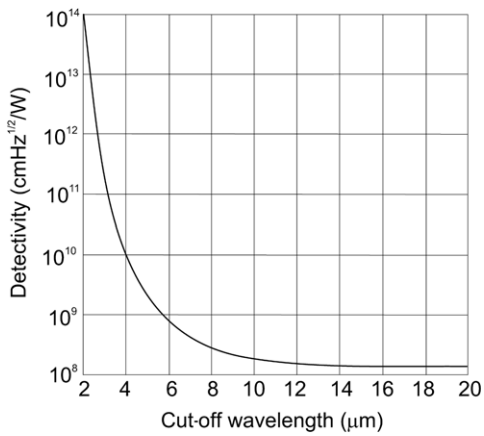


Figure 48. Detectivity needed to obtain NEDT = 0.1 K in a photon counter detector thermal imager as a function of cutoff wavelength (after [140]).

It is interesting to consider the performance requirements of near room temperature photodetectors for thermal cameras. The thermal resolution of IR thermal systems is usually characterized by NETD. It can be shown, that [142]

$$\mathrm{NETD} = \frac{4F^2\Delta f^{1/2}}{A_{\mathrm{d}}^{1/2} t_{\mathrm{op}}}\left[\int_{\lambda_a}^{\lambda_b} \frac{\mathrm{d}M}{\mathrm{d}T} D^*(\lambda)\,\mathrm{d}\lambda\right]^{-1}, \tag{46}$$

where F is the optics f-number, Δf is the frequency band, A_{d} is the detector area, t_{op} is the optics transmission and M is the spectral emittance of the blackbody described by the Planck's law.

As equation (46) shows, the thermal resolution improves with an increase in detector area. Increasing detector area results in reduced spatial resolution, however. Hence, a reasonable compromise between the requirement of high thermal and spatial resolution is necessary. Improvement of thermal resolution without spatial resolution worsening may be achieved by:

- increase in detector area combined with a corresponding increase in focal length and the objective aperture,
- improved detector performance,
- increase in the number of detectors.

Increase of aperture is undesirable because it increases the size, mass and price of an IR system. It is more proper to use a detector with higher detectivity. Another possibility is the application of multi-elemental sensor, which reduces each element bandwidth proportionally to the number of elements for the same frame rate and other parameters.

Figure 48 shows the dependence of detectivity on cutoff wavelength for a photon counter detector thermal imager with a resolution of 0.1 K. Detectivities of 1.9×10^8 cm Hz$^{1/2}$ W^{-1}, 2.3×10^8 cm Hz$^{1/2}$ W^{-1} and 2×10^9 cm Hz$^{1/2}$ W^{-1} are necessary to obtain NEDT = 0.1 K for 10 μm, 9 μm and 5 μm cutoff wavelength photon counter detectors, respectively. The above estimations indicate, that the ultimate performance of HgCdTe uncooled photodetectors is not sufficient to achieve the required 0.1 K thermal resolution. Thermal resolution below 0.1 K is achieved for starring thermal imagers containing thermal detector FPAs.

It is of interest to compare the performance of uncooled photon and thermal detectors in the MWIR ($\lambda = 5\,\mu$m) and LWIR ($\lambda = 10\,\mu$m) spectral range. In this comparison we follow

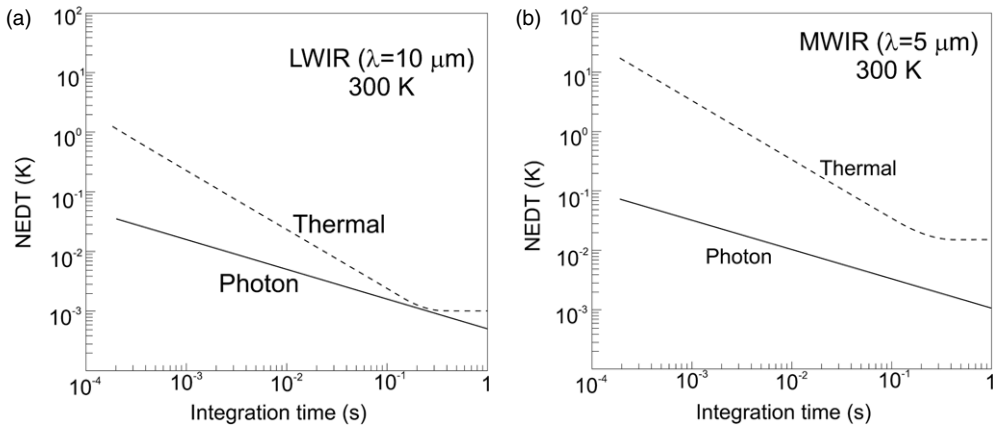


Figure 49. Theoretical NEDT comparison of uncooled thermal and HgCdTe uncooled photon LWIR (*a*) MWIR (*b*) detectors (after [22]).

the paper published by Kinch [22]. Figure 49 compares the theoretical NEDT of detectors operated at 290 K for $f/1$ optics and a 1 mil pixel size. As a photon detector N^+-π-P^+ HgCdTe photodiode is chosen, first proposed by Ashley and Elliott [143]. π designates an intrinsic region containing a p-type background dopant equal to 5×10^{14} cm^{-3} with a carrier lifetime limited by the Auger 7 process. It is also assumed that the detector node capacity can store the integrated charge due to detector dark current. Figure 49 shows that the ultimate performance of the uncooled HgCdTe photon detectors is far superior to the thermal detectors at wide frame rates and spectral bands.

Comparing both curves of figure 49 for thermal detectors we see, that for long integration times in the LWIR region excellent performance is achieved, with NEDT values below 10 mK for frame rates of 30 Hz. However, for snapshot systems with integration time below 2 ms, the available NEDT is above 100 mK even at the LWIR region. For the MWIR band the thermal detector has obvious performance limitations at any frame rate.

There are some ways of improving the performance of the photodetectors without cooling. A moderate p-type doping of the absorber detector region is widely used for some suppression of the Auger mechanisms [144–146]. More efficient suppression can be obtained with non-equilibrium depletion of the semiconductor [147]. However, the non-equilibrium mode devices suffer from a high level of flicker noise that make them useless for most of the practical applications that require detection of IR radiation in the low and moderate frequency range. An example is thermal imaging. Hence, non-equilibrium mode devices are omitted in our considerations.

A possible way to improve the performance of IR photodetectors is to reduce the physical volume of the semiconductor, thus reducing the amount of thermal generation (see section 2.5) [147, 148]. This must be done without decrease in quantum efficiency, optical area and FOV of the detector.

Thickness of the active region can be significantly reduced by enhanced absorption of radiation. This can be achieved by using multiple pass of radiation with a backside reflector. Even more efficient is the use of interference phenomena to set up a resonant cavity within the photodetector.

The other way to improve the performance of an IR photodetector is to increase the apparent 'optical' size of the detector in comparison with the actual physical size using a suitable concentrator which compresses impinging IR radiation. This must be achieved without

reduction of the acceptance angle, or at least, with limited reduction to angles required for fast optics of IR systems. Various types of suitable optical concentrators can be used, including optical cones and conical fibres. An efficient possible way to achieve an effective concentration of radiation is to immerse the photodetector in the hemispherical or hyperhemispherical lenses. The problems of matching the detectors to immersion lenses have been solved by the use of monolithic technology [144]. The technology is based on the epitaxy of HgCdZnTe on CdZnTe substrate. The HgCdZnTe serves as the sensitive element, while the immersion lens is formed directly in the transparent CdZnTe ($n = 2.7$) or GaAs ($n = 3.4$) substrates.

Initial efforts to produce near room temperature LWIR HgCdTe detectors have concentrated mainly on optimization of photoconductors, PEM detectors and Dember effect detectors (see table 9) [144–146]. Later, they were replaced with photovoltaic detectors which do not require electric or magnetic bias.

The room temperature photovoltaic device of conventional design suffers from

- poor quantum efficiency, and
- low differential resistance.

The problems of poor quantum efficiency and large series resistance have been solved through the adoption of sophisticated heterojunction architectures of photovoltaic devices in combination with the methods of reduction of physical size of the active element [149–155].

The practical realization of the multi-heterojunction device which consisted of a structure based on backside illuminated n^+-p-P photodiodes has been presented in several papers [150–154]. The individual detector elements were prepared by a combination of conventional dry etching, angled ion milling and angled thermal evaporation for contact metal deposition (see figure 50). A p-type active layer with a thickness of approximately 3 μm and doping level of about 10^{16} cm^{-3} was grown on GaAs substrates using MOCVD and *in situ* As doping. The delineation trenches in the epilayer were wet chemical etched using Br/ethylene glycol or Br/HBr solutions. The etch was followed by shallow ion beam milling using a Kaufman-type ion gun. The sample was placed at $\approx 45^\circ$ in the direction of the Ar^+ beam so that only one wall of the trench was exposed to the beam. The ion beam milling results in n-type conversion and the formation of n^+-p junctions on the uppermost surface and on one side of the trench walls. Cr/Au metallization was required to provide external contact to the device and to the short-circuited n^+-p junction formed at the base of the trench, effectively connecting side-by-side diodes in series. As a final step, the devices were passivated by thermal evaporation of 200 nm of CdZnTe followed by 300 nm of ZnS. For an ≈ 3 μm thick epilayer, the minimum practical multi-junction period using existing wet etching technology is ≈ 10 μm.

Figure 51 shows the room temperature spectral response of HgCdTe multi-heterojunction devices monolithically immersed to the CdZnTe lens. Generally, these room temperature devices have responsivities that are comparable to, or better than, photoelectromagnetic devices operating under the same conditions. Peltier cooled devices exhibit performance that is comparable to photoconductors operating at the same wavelength and temperatures. However, in contrast to photoconductors, multi-heterojunction detectors can be used at both low and very high frequencies. Heterodyne experiments indicate that the response time of LWIR devices at a wavelength of 10.6 μm is only of about 1 ns.

Multi-heterojunction HgCdTe devices are manufactured by Vigo System [155] for any wavelength within the 2–16 μm range, with active area sizes from a few micrometres to a few millimetres. Linear arrays up to 120 elements and small 2D arrays are manufactured as custom devices.

Optical immersion has been used almost exclusively for single element devices. The use of a single immersion lens to a large array is problematic in view of optical aberrations and

Table 9. HgCdTe uncooled photodetectors (after [141]).

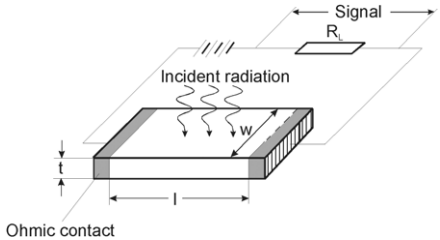
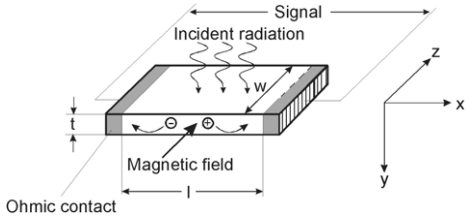
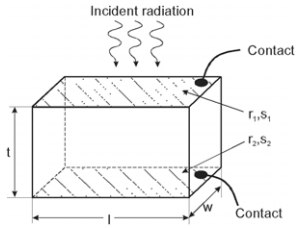
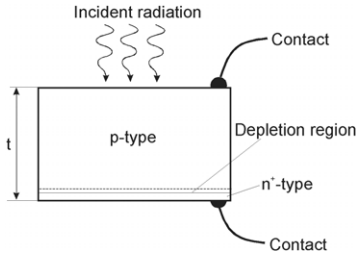
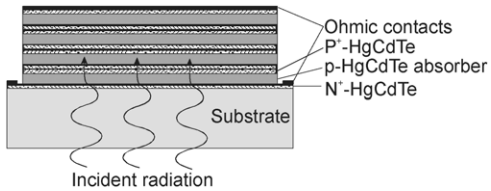
Mode of operation	Schematic of detector	Advantages	Disadvantages	Performance (optically immersed, $\lambda = 10.6\,\mu$m)
Photoconductor		• Low cost technology • High responsivity	• Bias required	At 300 K: • D^* up to 2×10^8 cm Hz$^{1/2}$ W^{-1} • Response time $\leqslant$1 ns At 220 K: • D^* up to 3×10^9 cm Hz$^{1/2}$ W^{-1} • Response time $\leqslant$10 ns
PEM detector		• No bias required • No flicker noise • Very short response time	• Bulky • Low resistance	At 300 K: • D^* up to 1×10^8 cm Hz$^{1/2}$ W^{-1} • Response time $\leqslant$10 ns
Dember detector		• No bias required • No flicker noise	• Low resistance • Require low resistance contacts	At 300 K: • D^* up to 1.5×10^7 cm Hz$^{1/2}$ W^{-1} • Response time $\leqslant$1 ns

Table 9. (Continued.)

Mode of operation	Schematic of detector	Advantages	Disadvantages	Performance (optically immersed, $\lambda = 10.6\,\mu$m)
Photodiode		• Fast response • No flicker noise	• Very low dynamic resistance • Low quantum efficiency • Influence of series resistance	At 220 K (multielement array): • D^* up to 2×10^8 cm Hz$^{1/2}$ W^{-1} • Response time $\leqslant 3$ ns
Stacked multiheterojunction photodiode		• No flicker noise • Extremely fast response • Can be used from DC to very high frequencies	• High dynamic resistance • Good quantum efficiency	• Generation–recombination limit of performance can be achieved in practice at any wavelength • Practical implementation requires well established epitaxial technology

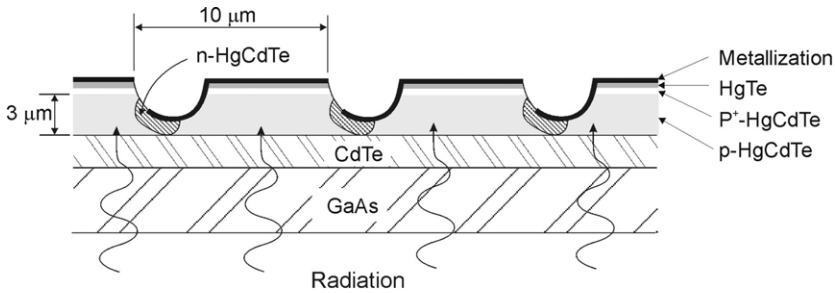


Figure 50. Backside illuminated multiple heterojunction device (after [154]).

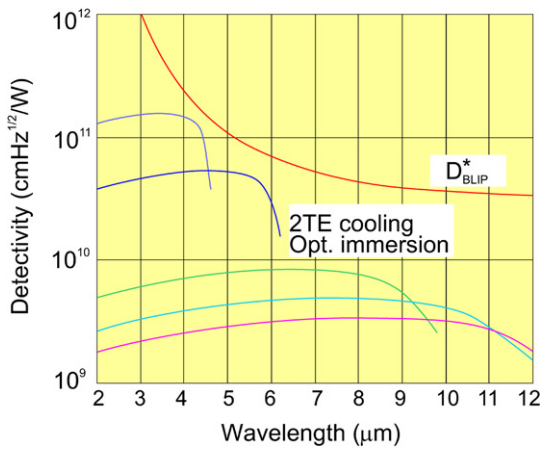


Figure 51. Measured detectivity of multi-heterojunction uncooled HgCdTe detectors. Optical immersion has been used to improve the performance (after [154]).

the large lens size. The problem can be solved by implementation of a small size 2D array monolithically integrated with microlenses. The small size of active element ($\approx 7 \times 7\,\mu$m) is beneficial for a good collection of photogenerated charge carriers especially in LWIR devices. The individual elements can be accessed individually or connected in series.

The devices are especially promising as uncooled 7.8–9.5 μm detectors that can be used for thermal imagers. Initial results are encouraging, indicating the potential for achieving $\approx 10^9\,\mathrm{cm\,Hz^{1/2}\,W^{-1}}$ at $\lambda \approx 9\,\mu$m. This would enable thermal resolution better than 0.1 K for starring thermal imagers operating with $f/1$ optics. Moreover, the devices exhibit very fast responses with a time constant of $\approx$1 ns. This is due to the short time required for photogenerated carriers in small size active region to reach heavily doped contacts and a short RC time constant as well.

In multiple heterojunction devices described previously (see figure 50), the junction's planes are perpendicular to the surface. More promising are the stacked multijunction photodiodes monolithically connected in series, which is shown in table 9. They are capable of achieving both good quantum efficiency and a large differential resistance. Each cell is composed of p-type doped narrow gap absorber and heavily doped N^+ and P^+ heterojunction contacts. The incoming radiation is absorbed only in absorber regions, while the heterojunction contacts collect the photogenerated charge carriers. Detectivity achieves a maximum for

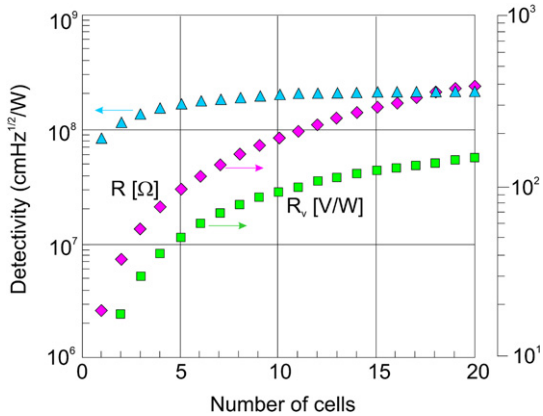


Figure 52. The parameters of multiple heterojunction photovoltaic devices (detectivity, resistance—R and voltage responsivity—R_v) as a function of the number of cells (after [156]).

15 cells. For practical reasons, the number of cells could be reduced to ≈10, with detectivity more than doubled in comparison to a single cell device (see figure 52). The device can be supplied with backside reflectors for double pass of IR radiation, optical resonant cavities and optical concentrators, such as immersion lenses. For a 10.6 μm device containing 5 cells and operated at room temperature, detectivity over to 10^9 cm Hz$^{1/2}$ W^{-1} can be achieved [156]. Practical implementation of the device would require well established low temperature epitaxial growth (MBE or MOCVD).

10. Summary

Over 40 years the HgCdTe ternary alloy has successfully fought off major challenges from alternative materials, but despite that it has more competitors today than ever before. These include InSb-based III–V materials, AlGaAs multiple quantum wells, InAs/GaInSb strain layer SLs and especially two types of thermal detectors: pyroelectric detectors and silicon bolometers. It is interesting to note that in practice, however, none of these competitors can compete in terms of fundamental properties. They may promise to be more manufacturable, but never to provide higher performance or, with the exception of thermal detectors, to operate at higher or even comparable temperatures.

The future applications of IR detector systems require

- higher pixel sensitivity,
- further increase in pixel density to above 10^6 pixels,
- cost reduction in IR imaging array systems due to less cooling sensor technology combined with the integration of detectors and signal processing functions (with much more on-chip signal processing),
- improvement in the functionality of IR imaging arrays through the development of multispectral sensors.

It is predicted that HgCdTe technology will continue, in the future, to expand the range of its capabilities because of its excellent properties.

Despite serious competition from alternative technologies and slower progress than expected, HgCdTe is unlikely to be seriously challenged for high-performance applications—applications requiring multispectral capability and fast response. However, non-uniformity

is a serious problem in the case of LWIR HgCdTe detectors. For applications that require operation in the LWIR band as well as two-colour MWIR/LWIR/VLWIR bands, HgCdTe will most probably not be the optimal solution. A type-II InAs/GaInSb SL structure is a relatively new alternative IR material system and has great potential for LWIR/VLWIR spectral ranges with performance comparable to HgCdTe with the same cutoff wavelength.

References

[1] Lawson W D, Nielson S, Putley E H and Young A S 1959 Preparation and properties of HgTe and mixed crystals of HgTe-CdTe *J. Phys. Chem. Solids* **9** 325–9

[2] Reine M B 2004 Fundamental properties of mercury cadmium telluride *Encyclopedia of Modern Optics* (London: Academic)

[3] Rogalski A, Adamiec K and Rutkowski J 2000 *Narrow-Gap Semiconductor Photodiodes* (Bellingham: SPIE Optical Engineering Press)

[4] Hansen G L, Schmit J L and Casselman T N 1982 Energy gap versus alloy composition and temperature in $Hg_{1-x}Cd_xTe$ *J. Appl. Phys.* **53** 7099–101

[5] Hansen G L and Schmit J L 1983 Calculation of intrinsic carrier concentration in $Hg_{1-x}Cd_xTe$ *J. Appl. Phys.* **54** 1639–40

[6] Weiler M H 1981 Magnetooptical properties of $Hg_{1-x}Cd_xTe$ alloys *Semiconductors and Semimetals* vol 16, ed R K Willardson and A C Beer (New York: Academic) pp 119–91

[7] Capper P P (ed) 1994 *Properties of Narrow Gap Cadmium-based Compounds* (*EMIS Datareviews Series No* 10) (London: IEE)

[8] Kim J S, Lowney J R and Thurber W R 1997 Transport properties of narrow-gap II–VI compound semiconductors *Narrow-gap II–VI Compounds for Optoelectronic and Electromagnetic Applications* ed P Capper (London: Chapman and Hall) pp 180–210

[9] Yadava R D S, Gupta A K and Warrier A V R 1994 Hole scattering mechanisms in $Hg_{1-x}Cd_xTe$ *J. Electron. Mater.* **23** 1359–78

[10] Rosbeck J P, Star R E, Price S L and Riley K J 1982 Background and temperature dependent current–voltage characteristics of $Hg_{1-x}Cd_xTe$ photodiodes *J. Appl. Phys.* **53** 6430–40

[11] Higgins W M, Seiler G N, Roy R G and Lancaster R A 1989 Standard relationships in the properties of $Hg_{1-x}Cd_xTe$ *J. Vac. Sci. Technol.* A **7** 271–5

[12] Dennis P N J, Elliott C T and Jones C L 1982 A method for routine characterization of the hole concentration in p-type cadmium mercury telluride *Infrared Phys.* **22** 167–9

[13] Amirtharaj P M and Burnett J H 1997 Optical properties of MCT *Narrow-Gap II–VI Compounds for Optoelectronic and Electromagnetic Applications* ed P Capper (London: Chapman and Hall) pp 133–79

[14] Finkman E and Schacham S E 1984 The exponential optical absorption band tail of $Hg_{1-x}Cd_xTe$ *J. Appl. Phys.* **56** 2896–900

[15] Chu J, Li B, Liu K and Tang D 1994 Empirical rule of intrinsic absorption spectroscopy in $Hg_{1-x}Cd_xTe$ *J. Appl. Phys.* **75** 1234–5

[16] Dornhaus R, Nimtz G and Schlicht B 1983 *Narrow-Gap Semiconductors* (Berlin: Springer)

[17] Lopes V C, Syllaios A J and Chen M C 1993 Minority carrier lifetime in mercury cadmium telluride *Semicond. Sci. Technol.* **8** 824–41

[18] Jóźwikowski K and Rogalski A 2000 Effect of dislocations on performance of LWIR HgCdTe photodiodes *J. Electron. Mater.* **29** 736–41

[19] Chen M C, Colombo L, Dodge J A and Tregilgas J H 1995 The minority carrier lifetime in doped and undoped p-type $Hg_{0.78}Cd_{0.22}Te$ liquid phase epitaxy films *J. Electron. Mater.* **24** 539–44

[20] Destefanis G and Chamonal J P 1993 Large improvement in HgCdTe photovoltaic detector performances at LETI *J. Electron. Mater.* **22** 1027–32

[21] Littler C L, Maldonado E, Song X N, You Z, Elkind J L, Seiler D G and Lowney J R 1991 Investigation of mercury interstitials in $Hg_{1-x}Cd_xTe$ alloys using resonant impact-ionization spectroscopy *J. Vac. Sci. Technol.* B **9** 1466–70

[22] Kinch M A 2000 Fundamental physics of infrared detector materials *J. Electron. Mater.* **29** 809–17

[23] Humpreys R G 1983 Radiative lifetime in semiconductors for infrared detectors *Infrared Phys.* **23** 171–5
Humpreys R G 1986 Radiative lifetime in semiconductors for infrared detectors *Infrared Phys.* **26** 337–42

[24] Elliott T, Gordon N T and White A M 1999 Towards background-limited, room-temperature, infrared photon detectors in the 3–13 μm wavelength range *Appl. Phys. Lett.* **74** 2881–3

[25] Casselman T N and Petersen P E 1980 A comparison of the dominant Auger transitions in p-type (Hg,Cd)Te *Solid State Commun.* **33** 615–19
[26] Casselman T N 1981 Calculation of the Auger lifetime in p-type $Hg_{1-x}Cd_xTe$ *J. Appl. Phys.* **52** 848–54
[27] White A M 1985 The characteristics of minority-carrier exclusion in narrow direct gap semiconductors *Infrared Phys.* **25** 729–41
[28] Beattie A and Landsberg P T 1959 Auger effect in semiconductors *Proc. R. Soc. Lond.* A **249** 16–29
[29] Krishnamurthy S and Casselman T N 2000 A detailed calculation of the Auger lifetime in p-type HgCdTe *J. Electron. Mater.* **29** 828–31
[30] Beattie R and White A M 1996 An analytic approximation with a wide range of applicability for electron initiated Auger transitions in narrow-gap semiconductors *J. Appl. Phys.* **79** 802–13
[31] Ciesla C M, Murdin B N, Phillips T J, White A M, Beattie A R, Langerak C J G M, Elliott C T, Pidgeon C R and Sivananthan S 1997 Auger recombination dynamics of $Hg_{0.795}Cd_{0.205}Te$ in the high excitation regime *Appl. Phys. Lett.* **71** 491–3
[32] Rose A 1963 *Concepts in Photoconductivity and Allied Problems* (New York: Interscience)
[33] Piotrowski J and Gawron W 1997 Ultimate performance of infrared photodetectors and figure of merit of detector material *Infrared Phys. Technol.* **38** 63–8
[34] Rogalski A 2003 Quantum well photoconductors in infrared detectors technology *J. Appl. Phys.* **93** 4355–91
[35] Smith D L and Mailhiot C 1987 Proposal for strained type II superlattice infrared detectors *J. Appl. Phys.* **62** 2545–8
[36] Grein C H, Young P M and Ehrenreich H 1992 Minority carrier lifetimes in ideal InGaSb/InAs superlattice *Appl. Phys. Lett.* **61** 2905–7
[37] Youngdale E R, Meyer J R, Hoffman C A, Bartoli F J, Grein C H, Young P M, Ehrenreich H, Miles R H and Chow D H 1994 Auger lifetime enhancement in InAs–$Ga_{1-x}In_xSb$ superlattices *Appl. Phys. Lett.* **64** 3160–2
[38] Grein C H, Cruz H, Flatte M E and Ehrenreich H 1994 Theoretical performance of very long wavelength InAs/$In_xGa_{1-x}Sb$ superlattice based infrared detectors *Appl. Phys. Lett.* **65** 2530–2
[39] Grein C H, Young P M, Flatté M E and Ehrenreich H 1995 Long wavelength InAs/InGaSb infrared detectors: optimization of carrier lifetimes *J. Appl. Phys.* **78** 7143–52
[40] Kruse P W, McGlauchlin L D and McQuistan R B 1962 *Elements of Infrared Technology* (New York: Wiley)
[41] Reine M B, Sood A K and Tredwell T J 1981 Photovoltaic infrared detectors *Semiconductors and Semimetals* vol 18, ed R K Willardson and A C Beer (New York: Academic) pp 201–311
[42] Rogalski A 1988 Analysis of the R_0A product in n^+–p $Hg_{1-x}Cd_xTe$ photodiodes *Infrared Phys.* **28** 139–53
[43] Reine M B 2000 Photovoltaic detectors in HgCdTe *Infrared Detectors and Emitters: Materials and Devices* ed P Capper and C T Elliott (London: Chapman and Hall) pp 313–76
[44] Rogalski A and Larkowski W 1985 Comparison of photodiodes for the 3–5.5 μm and 8–12 μm spectral regions *Electron Technol.* **18** 55–69
[45] Burstein E, Davisson J W, Bell E E, Turner W J and Lipson H G 1954 *Phys. Rev.* **93** 65
[46] Borrello S and Levinstein H 1962 Preparation and properties of mercury-doped germanium *J. Appl. Phys.* **33** 2947–50
[47] Long D and Schmit J L 1970 Mercury–cadmium telluride and closely related alloys *Semiconductors and Semimetals* vol 5, ed R K Willardson and A C Beer (New York: Academic) pp 175–255
[48] Norton P 2002 HgCdTe infrared detectors *Opto-Electron. Rev.* **10** 159–74
[49] Verie C and Granger R 1965 Propriétés de junctions p-n d'alliages $Cd_xHg_{1-x}Te$ *C. R. Acad. Sci.* **261** 3349–52
[50] Verie G C and Sirieix M 1972 Gigahertz cutoff frequency capabilities of CdHgTe photovoltaic detectors at 10.6 μm *IEEE J. Quant. Electron.* **8** 180–4
[51] Bartlett B E, Charlton D E, Dunn W E, Ellen P C, Jenner M D and Jervis M H 1969 Background limited photoconductive detectors for use in the 8–14 micron atmospheric window *Infrared Phys.* **9** 35–6
[52] Kinch M A, Borrello S R and Simmons A 1977 0.1 eV HgCdTe photoconductive detector performance *Infrared Phys.* **17** 127–35
[53] Elliott C T, Day D and Wilson B J 1982 An integrating detector for serial scan thermal imaging *Infrared Phys.* **22** 31–42
[54] Blackburn A, Blackman M V, Charlton D E, Dunn W A E, Jenner M D, Oliver K J and Wotherspoon J T M 1982 The practical realization and performance of SPRITE detectors *Infrared Phys.* **22** 57–64
[55] Boyle W S and Smith G E 1970 Charge-coupled smiconductor devices *Bell Syst. Tech. J.* **49** 587–93
[56] Thom R High density infrared detector arrays *US Patent No* 4,039,833, 8/2/77
[57] Baker I M and Ballinga R A 1984 Photovoltaic CdHgTe-silicon hybrid focal planes *Proc. SPIE* **510** 121–9
[58] Norton P 2003 Detector focal plane array technology *Encyclopedia of Optical Engineering* ed R Driggers (New York: Dekker) pp 320–48

[59] Kinch M A 1994 MIS devices in HgCdTe *Properties of Narrow Gap Cadmium-based Compounds* (*EMIS Datareviews Series No* 10) ed P Capper (London: IEE) pp 359–63
[60] Reine M B 2000 Photovoltaic detectors in HgCdTe *Infrared Detectors and Emitters: Materials and Devices* ed P Capper and C T Elliott (London: Chapman and Hall) pp 313–76
[61] Yu T C and Brebrick R F 1994 Phase diagrams for HgCdTe *Properties of Narrow Gap Cadmium-Based Compounds* (*EMIS Datareviews Series No* 10) ed P P Capper (London: IEE) pp 55–63
[62] Colombo L, Chang R R, Chang C J and Baird B A 1988 Growth of Hg-based alloys by the travelling heater method *J. Vac. Sci. Technol.* A **6** 2795–9
[63] Gertner E R 1985 Epitaxial mercury cadmium telluride *Ann. Rev. Mater. Sci.* **15** 303–28
[64] Tung T, DeArmond L V, Herald R F, Herning P E, Kalisher M H, Olson D A, Risser R F, Stevens A P and Tighe S J 1992 State of the art of Hg-melt LPE HgCdTe at Santa Barbara Research Center *Proc. SPIE* **1735** 109–31
[65] Irvine S J C 1997 Metal-organic vapour phase epitaxy *Narrow-gap II–VI Compounds for Optoelectronic and Electromagnetic Applications* ed P Capper (London: Chapman and Hall) pp 71–96
[66] Wu O K, deLyon T J, Rajavel R D and Jensen J E 1997 Molecular beam epitaxy of HgCdTe *Narrow-gap II–VI Compounds for Optoelectronic and Electromagnetic Applications* ed P Capper (London: Chapman and Hall) pp 97–130
[67] Wijewarnasuriya P S *et al* 1998 MBE p-on-n $Hg_{1-x}Cd_xTe$ heterostructure detectors on silicon substrates *J. Electron. Mater.* **27** 546–9
[68] DeLyon T J, Jensen J E, Gorwitz M D, Cockrum C A, Johnson S M and Venzor G M 1999 MBE growth of HgCdTe on silicon substrates for large-area infrared focal plane arrays: a review of recent progress *J. Electron. Mater.* **28** 705–11
[69] Johnson S M *et al* 2004 HgCdTe/Si materials for long wavelength infrared detectors *J. Electron. Mater.* **33** 526–30
[70] Carmody M *et al* 2004 Long wavelength infrared, molecular beam epitaxy, HgCdTe-on-Si diode performance *J. Electron. Mater.* **33** 531–7
[71] Turner A *et al* 1994 Producibility of VIP^{TM} scanning focal plane arrays *Proc. SPIE* **2228** 237–48
[72] Baker I M 1997 Photovoltaic IR detectors *Narrow-gap II–VI Compounds for Optoelectronic and Electromagnetic Applications* ed P Capper (London: Chapman and Hall) pp 450–73
[73] Tribolet P, Chatard J P, Costa P and Manissadjian A 1998 Progress in HgCdTe homojunction infrared detectors *J. Cryst. Growth* **184/185** 1262–71
[74] Bajaj J 1998 State-of-the-art HgCdTe materials and devices for infrared imaging *Physics of Semiconductor Devices* ed V Kumar and S K Agarwal (New Delhi: Narosa Publishing House) pp 1297–309
[75] Varavin V S, Vasiliev V V, Dvoretsky S A, Mikhailov N N, Ovsyuk V N, Sidorov Yu G, Suslyakov A O, Yakushev M V and Aseev A L 2003 HgCdTe epilayers on GaAs: growth and devices *Opto-Electron. Rev.* **11** 99–111
[76] Pultz G N, Norton P W, Krueger E E and Reine M B 1991 Growth and characterization of p-on-n HgCdTe liquid-phase epitaxy heterojunction material for 11–18 μm applications *J. Vac. Sci. Technol.* B **9** 1724–30
[77] Mitra P, Case F C and Reine M B 1998 Progress in MOVPE of HgCdTe for advanced infrared detectors *J. Electron. Mater.* **27** 510–20
[78] Tung T, Kalisher M H, Stevens A P and Herning P E 1987 Liquid-phase epitaxy of $Hg_{1-x}Cd_xTe$ from Hg solution: a route to infrared detector structures *Mat. Res. Symp. Proc.* **90** 321–56
[79] T J DeLyon, Jensen J E, Gorwitz M D, Cockrum C A, Johnson S M and Venzor G M 1999 MBE growth of HgCdTe on silicon substrates for large-area infrared focal plane arrays: a review of recent progress *J. Electron. Mater.* **28** 705–11
[80] Ameurlaire J F and Cohen-Solal G D 1974 *US Patent No* 3,845,494
[81] Cohen-Solal G D and Lussereau A G 1976 *US Patent No* 4,998,774
[82] Norton P 1998 Status of infrared detectors *Proc. SPIE* **3379** 102–14
[83] Bajaj J 2000 State-of-the-art HgCdTe infrared devices *Proc. SPIE* **3948** 42–54
[84] Reine M B, Tobin S P, Norton P W and LoVecchio P 2004 Very long wavelength (>15 μm) HgCdTe photodiodes by liquid phase epitaxy *Proc. SPIE* **5564** 54–64
[85] Bubulac L O *et al* 1997 High performance SWIR HgCdTe detector arrays *J. Electron. Mater.* **26** 649–55
[86] Kozlowski L J, Vural K, Arias J M, Tennant W E and DeWames R E 1997 Performance of HgCdTe, InGaAs and quantum well GaAs/AlGaAs staring infrared focal plane arrays *Proc. SPIE* **3182** 2–13
[87] DeWames R E, Edwall D D, Zandian M, Bubulac L O, Pasko J G, Tennant W E, Arias J M and D'Souza A 1998 Dark current generating mechanisms in short wavelength infrared photovolatic detectors *J. Electron. Mater.* **27** 722–6

[88] Tennant W E, Cabelli S and Spariosu K 1999 Prospects of uncooled HgCdTe detector technology *J. Electron. Mater.* **28** 582–8
[89] Mitra P, Case F C, Reine M B, Parodos T, Tobin S P and Norton P W 1999 MOVPE growth of HgCdTe for high performance 3–5 μ photodiodes operating at 100–180 K *J. Electron. Mater.* **28** 589–95
[90] Rogalski A and Ciupa R 1999 Performance limitation of short wavelength infrared InGaAs and HgCdTe photodiodes *J. Electron. Mater.* **28** 630–6
[91] Chen M C, List R S, Chandra D, Bevan M J, Colombo L and Schaake H F 1996 Key performance-limiting defects in p-on-n HgCdTe heterojunction infrared photodiodes *J. Electron. Mater.* **25** 1375–82
[92] Kinch M 2001 HDVIP^{TM} FPA technology at DRS *Proc. SPIE* **4369** 566–78
[93] Johnson S M, Rhiger D R, Rosbeck J P, Peterson J M, Taylor S M and Boyd M E 1992 Effect of dislocations on the electrical and optical properties of long-wavelength infrared HgCdTe photovoltaic detectors *J. Vac. Sci. Technol.* B **10** 1499–506
[94] Vural K *et al* 1999 2048 × 2048 HgCdTe focal plane arrays for astronomy applications *Proc. SPIE* **3698** 24–35
[95] http://www.rsc.rockwell.com/imaging/
[96] http://compoundsemiconductor.net/articles/news/6/8/5/1
[97] Reago D, Horn S, Campbell J and Vollmerhausen R 1999 Third generation imaging sensor system concepts *Proc. SPIE* **3701** 108–17
[98] Norton P, Campbell J, Horn S and Reago D 2000 Third-generation infrared imagers *Proc. SPIE* **4130** 226–36
[99] Kozlowski L J and Kosonocky W F 1995 Infrared detector arrays *Handbook of Optics* ed M Bass *et al* (New York: McGraw-Hill) chapter 23
[100] Norton P R 1999 Infrared detectors in the next millennium *Proc. SPIE* **3698** 652–65
[101] Levine B F 1993 Quantum-well infrared photodetectors *J. Appl. Phys.* **74** R1–81
[102] Tidrow M Z *et al* 1999 Device physics and focal plane applications of QWIP and MCT *Opto-Electron. Rev.* **7** 283–96
[103] Sarusi G 2003 QWIP or other alternatives for third generation infrared systems *Infrared Phys. Technol.* **44** 439–44
[104] Horn S, Norton P, Cincotta T, Stolz A, Benson D, Perconti P and Campbell J 2003 Challenges for third-generation cooled imagers *Proc. SPIE* **5074** 44–51
[105] Reine M B, Hairston A, O'Dette P, Tobin S P, Smith F T J, Musicant B L, Mitra P and Case F C 1998 Simultaneous MW/LW dual-band MOCVD HgCdTe 64 × 64 FPAs *Proc. SPIE* **3379** 200–12
[106] Rajavel R D *et al* 1998 Molecular beam epitaxial growth and performance of integrated multispectral HgCdTe photodiodes for the detection of mid-wave infrared radiation *J. Cryst. Growth* **184** 1272–8
[107] Rogalski A 2000 Dual-band infrared detectors *J. Infrared Millim. Waves* **19** 241–58
[108] Kinch M A 2001 HDVIP^{TM} FPA technology at DRS *Proc. SPIE* **4369** 566–78
[109] Cabanski W, Breiter R, Mauk K-H, Rode W, Ziegler J, Schneider H, Walther M and Oelmaier R 2003 Status of 3rd gen focal plane array IR detection modules at AIM *Proc. SPIE* **5074** 72–82
[110] Smith E P G *et al* 2004 HgCdTe focal plane arrays for dual-color mid- and long-wavelength infrared detection *J. Electron. Mater.* **33** 509–16
[111] Ballet P *et al* 2004 Dual-band infrared detectors made on high-quality HgCdTe epilayers grown by molecular beam epitaxy on CdZnTe of CdTe/Ge substrates *J. Electron. Mater.* **33** 667–72
[112] Melngailis J and Harman T C 1970 Single-crystal lead-tin chalcogenides *Semiconductors and Semimetals* vol 5, ed R K Willardson and A C Beer (New York: Academic) pp 111–74
[113] Harman T C and Melngailis J 1974 Narrow gap semiconductors *Applied Solid State Science* vol 4, ed R Wolfe (New York: Academic) pp 1–94
[114] Gunapala S D *et al* 2001 Recent developments and applications of quantum well infrared photodetector focal plane arrays *Opto-Electron. Rev.* **8** 150–63
[115] Schneider H, Koidl P, Walther M, Fleissner J, Rehm R, Diwo E, Schwarz K and Weimann G 2001 Ten years of QWIP development at Fraunhofer *Infrared Phys. Technol.* **42** 283–9
[116] Gunapala S D, Bandara S V, Liu J K, Rafol B and Mumolo J M 2004 640 × 512 pixel long-wavelength infrared narrowband, multiband, and broadband QWIP focal plane arrays *IEEE Trans. Electron Devices* **50** 2353–60
[117] Rogalski A and Piotrowski J 1988 Intrinsic infrared detectors *Prog. Quant. Electron.* **12** 87–289
[118] Longo J T, Cheung D T, Andrews A M, Wang C C and Tracy J M 1978 Infrared focal planes in intrinsic semiconductors *IEEE Trans. Electron Devices* **ED–25** 213–32
[119] Zogg H 1998 Lead chalcogenide on silicon infrared sensor arrays *Opto-Electron. Rev.* **6** 37–46
[120] Zogg H and Ishida A 2001 IV–VI (lead chalcogenide) infrared sensors and lasers *Infrared Detectors and Emitters: Materials and Devices* ed P Capper and C T Elliott (Boston: Kluwer) pp 43–75
[121] Baars J 1982 New aspects of the material and device technology of intrinsic infrared photodetectors *Physics and Narrow Gap Semiconductors* ed E Gornik *et al* (Berlin: Springer) pp 280–2

[122] Love P J, Ando K J, Bornfreund R E, Corrales E, Mills R E, Cripe J R, Lum N A, Rosbeck J P and Smith M S 2002 Large-format infrared arrays for future space and ground-based astronomy applications *Proc. SPIE* **4486** 373–84
[123] Ashley T *et al* 2003 Large format MWIR focal plane arrays *Proc. SPIE* **4820** 400–5
[124] Zandian M *et al* 2003 Mid-wavelength infrared p-on-on $Hg_{1-x}Cd_xTe$ heterostructure detectors: 30–120 Kelvin state-of-the-art performance *J. Electron. Mater.* **32** 803–9
[125] Ettenberg M H, Lange M J, O'Grady M T, Vermaak J S, Cohen M J and Olsen G H 2000 A room temperature 640 × 512 pixel near-infrared InGaAs focal plane array *Proc. SPIE* **4028** 201–7
[126] Tidrow M Z 2000 Device physics and state-of-the-art of quantum well infrared photodetectors and arrays *Mater. Sci. Eng.* B **74** 45–51
[127] Schneider H, Koidl P, Walther M, Fleissner J, Rehm R, Diwo E, Schwarz K and Weimann G 2001 Ten years of QWIP development at Fraunhofer *Infrared Phys. Technol.* **42** 283–9
[128] Jhabvala M, Choi K, Goldberg A, La A and Gunapala S 2004 Development of a 1 k × 1 k GaAs QWIP far IR imaging array *Proc. SPIE* **5167** 175–85
[129] Gunapala S D *et al* 2003 640 × 512 pixel narrow-band, four-band, and broad-band quantum well infrared photodetector focal plane arrays *Infrared Phys. Technol.* **44** 411–25
[130] Bürkle L and Fuchs F 2002 InAs/(GaIn)Sb superlattices: a promising material system for infrared detection 2002 *Handbook of Infrared Detection and Technologies* ed M Henini and M Razeghi (Oxford: Elsevier) pp 159–89
[131] Yang O K, Pfahler C, Schmitz J, Pletschen W and Fuchs F 2003 Trap centers and minority carrier lifetimes in InAs/GaInSb superlattice long wavelength photodetectors *Proc. SPIE* **4999** 448–56
[132] Johnson J L 2000 The InAs/GaInSb strained layer superlattice as an infrared detector material: an overview *Proc. SPIE* **3948** 118–32
[133] Fuchs F, Bürkle L, Hamid R, Herres N, Pletschen W, Sah R E, Kiefer R and Schmitz J 2001 Optoelectronic properties of photodiodes for the mid- and far-infrared based on the InAs/GaSb/AlSb materials family *Proc. SPIE* **4288** 171–82
[134] Young M H, Chow D H, Hunter A T and Miles R H 1998 Recent advances in $Ga_{1-x}In_xSb$/InAs superlattice IR detector materials *Appl. Surf. Sci.* **123/124** 395–9
[135] Cabanski W A, Eberhardt K, Rode W, Wendler J C, Ziegler J, Fleissner J, Fuchs F, Rehm R, Schmitz J, Schneider H and Walther M 2004 Third generation focal plane array IR detection modules and applications *Proc. SPIE* **5406** 184–92
[136] Rogalski A 2001 Hg-based alternatives to MCT *Infrared Detectors and Emitters: Materials and Devices* ed P Capper and C T Elliott (Boston: Kluwer Academic) pp 377–400
[137] Kruse P W and Skatrud D D 1997 *Semiconductors and Semimetals* ed P W Kruse and D D Skatrud (San Diego: Academic)
[138] Kruse P W 1999 Uncooled IR focal plane arrays *Opto-Electron. Rev.* **7** 253–8
[139] Kruse P W 2001 *Uncooled Thermal Imaging. Arrays, Systems, and Applications* (Bellingham: SPIE Optical Engineering Press)
[140] Murphy D *et al* 2001 High sensitivity (25 μm pitch) microbolometer FPAs and application development *Proc. SPIE* **4369** 222–34
[141] Piotrowski J and Rogalski A 2004 Uncooled long wavelength infrared photon detectors *Infrared Phys. Technol.* **46** 115–31
[142] Rogalski A 2000 *Infrared Detectors* (Amsterdam: Gordon and Breach)
[143] Ashley T and Elliott C T 1985 Non-equilibrium devices for infrared detection *Electron. Lett.* **21** 451–2
[144] Piotrowski J, Galus W and Grudzień M 1991 Near room-temperature IR photodetectors *Infrared Phys.* **31** 1–48
[145] Piotrowski J 1995 $Hg_{1-x}Cd_xTe$ detectors *Infrared Photon Detectors* ed A Rogalski (Bellingham: SPIE Optical Engineering Press) pp 391–493
[146] Piotrowski J and Rogalski A 1997 Photoelectromagnetic, magnetoconcentration and Dember infrared detectors *Narrow-Gap II–VI Compounds and Electromagnetic Applications* ed P Capper (London: Chapman and Hall) pp 506–25
[147] Elliott C T and Gordon N T 1993 Infrared detectors *Handbook on Semiconductors* vol 4, ed C Hilsum (Amsterdam: North-Holland) pp 841–936
[148] Elliott C T 1990 Non-equilibrium mode of operation of narrow-gap semiconductor devices *Semicond. Sci. Technol.* **5** S30–7
[149] Piotrowski J and Gawron W 1995 Extension of longwavelength IR photovoltaic detector operation to near room-temperatures *Infrared Phys. Technol.* **36** 1045–51
[150] Musca C, Antoszewski J, Dell J, Faraone L, Piotrowski J and Nowak Z 1998 Multi-heterojunction large area HgCdTe long wavelength infrared photovoltaic detector for operation at near room temperature *J. Electron. Mater.* **27** 740–6

[151] Piotrowski J, Nowak Z, Antoszewski J, Musca C, Dell J and Faraone L 1998 A novel multi-heterojunction HgCdTe long-wavelength infrared photovoltaic detector for operation under reduced cooling conditions *Semicond. Sci. Technol.* **13** 1209–14

[152] Gawron W and Rogalski A 2002 HgCdTe buried multi-junction photodiodes fabricated by the liquid phase epitaxy *Infrared Phys. Technol.* **43** 157–63

[153] Piotrowski J, Grudzień M, Nowak Z, Orman Z, Pawluczyk J, Romanis M and Gawron W 2000 Uncooled photovoltaic $Hg_{1-x}Cd_xTe$ LWIR detectors *Proc. SPIE* **4130** 175–84

[154] Piotrowski A, Madejczyk P, Gawron W, Kłos K, Pawluczyk J, Grudzień M, Piotrowski J and Rogalski A MOCVD HgCdTe heterostructures for uncooled infrared photodetectors *Proc. SPIE* **5732** 273–84

[155] Vigo System data sheets

[156] Piotrowski J, Brzozowski P and Jóźwikowski K 2003 Stacked multijunction photodetectors of long wavelength radiation *J. Electron. Mater.* **32** 672–6