Opto-Electronics Review 31 (2023) e144566

Opto-Electronics Review

journal homepage: https://journals.pan.pl/opelre

Cryogenic solutions for IR detectors – a guideline for selection

René Griot[1*], Christophe Vasse[1], Roel Arts[2], Ruslan Ivanov[3], Linda Höglund[3], Eric Costard[3]

[1]Thales LAS France, 4 rue Marcel Doret, 31700 Blagnac, France
[2]Thales Cryogenics bv, Hooge Zijde 14, 5626 DC Eindhoven, The Netherlands
[3]IRnova, Isafjordsgatan 26, SE-164 40 Kista, Sweden

Article info

Article history:
Received 21 Oct. 2022
Received in revised form 23 Dec. 2022
Accepted 04 Jan. 2023
Available on-line 27 Feb. 2023

Keywords:
Cryocoolers; Stirling rotary coolers; Stirling linear coolers; pulse tube coolers; infrared detectors.

Abstract

As long as high resolution or long-range observation is to be achieved using infrared detection, it will be necessary to cool down the detector in order to reach the best sensitivity and dynamics. This paper describes different cooling solutions currently used for this purpose discussing advantages and drawbacks. Some guideline is given for cooler choice and selection. The focus is on rotary Stirling coolers illustrated by description of the RMs1 cooler dedicated to high operating temperature size, weight, and power infrared detectors. A user case study is presented with cooler power consumption and cool down time of the RMs1 cooler when integrated in IRnova's Oden MW IDDCAs.

1. Introduction

Although some pre-WW2 examples can be found, applications for infrared (IR) detection appeared in the 1960s and since then many improvements have been made in terms of autonomy and portability, extending the scope of applications [1]. Photon IR detectors need to operate at low temperature and therefore have to be associated with a cooling device. This constraint remains today to reach the best sensitivity and dynamics. Anytime high resolution, high magnification, or long-range observation are at stake, the cooling function will be necessary. This paper presents classical cooling solutions used in thermal detection pointing out pros and cons for each system and giving newcomers some hints and guideline for the cooler selection and use.

2. How to cool

Photon detectors must be cooled to a temperature sufficiently low to reduce the dark current and thermal noise to a level where it no longer impedes imaging the infrared radiation. This low temperature is necessary to prevent the thermal generation of charge carriers which are transitions that parasitize the optical ones and make uncooled devices very noisy [1]. As it all started using liquid nitrogen cooling, the reference and historically used cooling temperature for IR detectors is 77 K. Progress made in detector design now allows to operate at higher temperatures, up to 150 K, thereby reducing cooling power requirements. By definition the term "cryogenic cooling" is used for cooling applications below 200 K.

The first cooling method in early laboratory experiments with IR detection was liquid nitrogen providing a stable and constant 77 K, as long as nitrogen in the liquid phase remained present before total evaporation. Given the availability of mass-produced liquid nitrogen, even in the early days of IR detection, this was and is a convenient way to operate in a lab environment but is inconvenient in most practical applications due to size, weight, and autonomy criteria.

The cooling temperature range for IR detection is roughly from 50 K to 200 K, depending on technology, wavelength range of interest, and sensitivity needed. Different types of coolers are currently used for that purpose: Stirling coolers, pulse-tube coolers, Joule-Thomson coolers, thermoelectric coolers. Each of these cooling technologies have their advantages, drawbacks, and limitations. There is no perfect fit, just compromise. The choice of cooling technology and the cooler will depend on specific operating conditions and requirements for each application.

*Corresponding author at: rene.griot@fr.thalesgroup.com

https://doi.org/10.24425/opelre.2023.144566
1896-3757/ Association of Polish Electrical Engineers (SEP) and Polish Academic of Sciences (PAS). Published by PAS

As there is a significant temperature gradient to the detector, managing the heat loss to environment is an important aspect of any infrared detection device. For this purpose, the component to be cooled will be encapsulated in a vacuum dewar surrounding the cooler cold finger and containing an infrared detector device. Main contribution to heat losses are: thermal conduction through the cold finger, gas conduction and convection in the residual vacuum, thermal conduction in focal plane wire bonding, ambient heat radiation that can be reduced using thermal screens. For a given cooling device, thermal losses will be part of the thermal load reducing available cooling capacity and increasing power consumption.

3. Cooling solutions

3.1. *Stirling coolers*

The Stirling cycle is a thermodynamic cycle derived from the motor created by Reverend Robert Stirling in 1816 to rival the steam engine converting heat flow into mechanical power. This cycle is reversible, and the reversed Stirling cycle is used for cooling applications converting mechanical power into a flow of heat (Fig. 1).

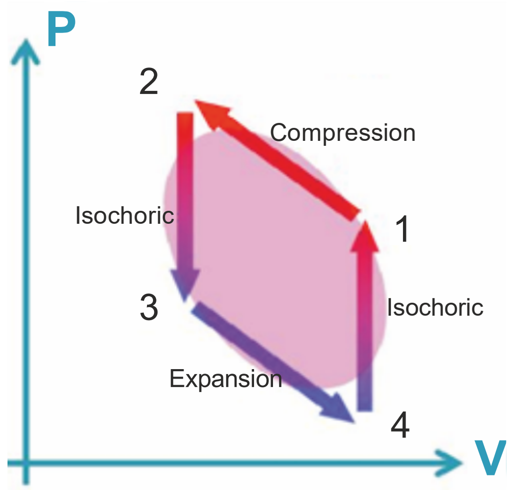

Fig. 1. Reversed Stirling cycle.

Four strokes are combined in the Stirling cycle: compression, expansion, and two isochoric transformations. Stirling coolers are mechanical machines built to apply the Stirling cycle to a working fluid. To do so, they combine two movements:

- reciprocating a motion of a compressor piston alternating compression and expansion phases of the gas present in the working chamber. This stands for two strokes of the reversed Stirling cycle.
- reciprocating a movement of a displacer allowing transfer of a gas volume from cold to hot side to perform two remaining isochoric strokes of the Stirling cycle. As the gas travels from the cold side to the hot side and back, the gas passes through a regenerator, a heat exchanger, to emphasize the cooling effect.

The compressor piston movement creates a sine pressure wave for the gas present in the working chamber centred on the static gas filling pressure. Gas present in the chamber is periodically heated up during the compression phase and cooled down in the expansion phase (Fig. 2).

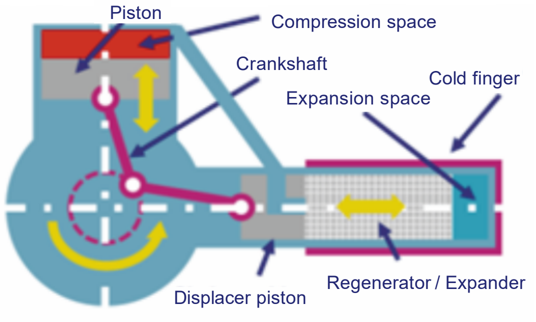

Fig. 2. Rotary Stirling cooler.

The two movements: compressor piston and displacer are tuned with a 90° phase shift to get a cooling effect at the tip of the cold finger (while there is heat rejection at the hot side). Due to the phase shift in the combination of compressor piston and displacer motion, a volume of gas will be present at the tip of the cold finger during the pressure release phase (cooling effect). This portion of gas is then rejected to the hot side due to a displacer movement, so no gas is present in the cold tip zone during the compression phase (heating effect). Gas transfer from the cold to the hot side is done through the regenerator, so cooled gas is heated up travelling from the cold to the hot side and cooled again on the way back to the cold end to benefit from an additional cooling effect. Cycle after cycle temperature drops at the cold end.

Stirling coolers are closed-circuit regenerative engines and use helium gas as the working fluid. There is no contact between helium and parts to be cooled and there is no phase transformation in this process. Such system can be miniaturized to fit size weight and power requirements of most of IR detection applications.

There are two main types of Stirling cryocoolers: linear and rotary.

Working principles are globally the same, the difference lies in the actuation mode of the compressor piston:

- in a linear cooler, the compressor piston is placed in a coil and actuated in a reciprocating movement by ways of electro-magnetism.
- in a rotary cooler, there is a motor with a rotating shaft, the compressor piston is actuated from the shaft using a conrod – crankshaft assembly.

Electrical supply for the linear cooler is AC current. The frequency for piston movement is fixed and the cooling power regulation is done by controlling the piston stroke (supply voltage).

Rotary coolers use brushless direct current (BLDC), 3-phase motors. The relative position of the rotor to the stator is monitored by the Hall effect sensors (3). An electronic driver is necessary to manage the voltage supply distribution from one winding to the other based on the signal from the Hall effect sensors. Piston stroke is fixed, and cooling power regulation is done by controlling the motor rotary speed (supply voltage).

Two cooler configurations can be distinguished:

- the split configuration: cooler is made of two main separate units, compressor and cold finger. Pressure wave is generated at compressor side and transmitted to cold finger through a metallic pipe. Displacer movement in the cold finger part is from pneumatic resonance tuned to provide the required 90° phase shift or by a separate actuator. Parts to be cooled are placed at the tip of the cold finger.

- the integral configuration: two main parts – compressor and cold finger, are integrated in the same compact unit. In general, both compressor piston and displacer are directly actuated from the main shaft, phase shift is fixed and guaranteed. Gas path from compressor to cold finger is reduced (a duct hole in cooler body) reducing pressure losses on the way.

Linear coolers (Fig. 3) are typically made in the split configuration.

Rotary coolers (Fig. 4) can be made both in split and integral configuration, with the integral configuration resulting in an additional advantage on efficiency as the displacer part is connected to the motor, securing phase shift value.

Most of the induced vibration produced by the cooler comes from the compressor (unbalancing, shocks). The split configuration is then a good way to reduce a vibration level at the cold tip.

The choice between linear or rotary cooler depends on application requirements. Rotary coolers are usually smaller, lighter, more efficient, and cheaper than linear. Linear coolers usually show higher level of reliability with lower noise and vibration levels.

Fig. 3. LSF9997 split linear Stirling cooler.

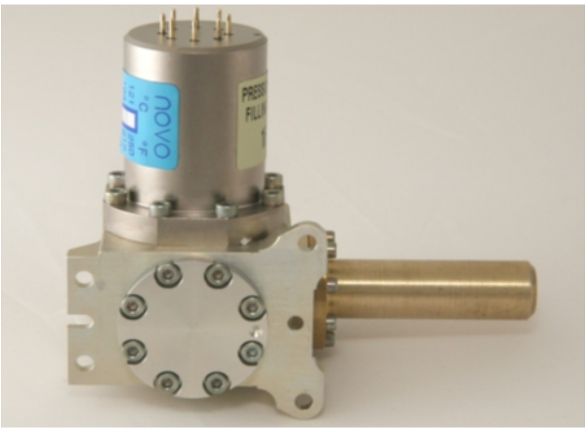

Fig. 4. RM2 integral rotary Stirling coolers.

3.2. Pulse-tube coolers

Pulse-tube coolers (Figs. 5 and 6) use the same kind of compressor unit as in the linear Stirling cooler to provide a pressure wave. The main difference is in the cold finger, where there is still a regenerator but no displacer. The working gas undergoes four stages quite similar to the Stirling cycle described in the previous section, but in the pulse-tube cold finger, there are no moving parts to move the working gas between the hot and cold sides. The geometry of the device itself allows the gas to flow to the cold side just before the pressure release (cooling effect)

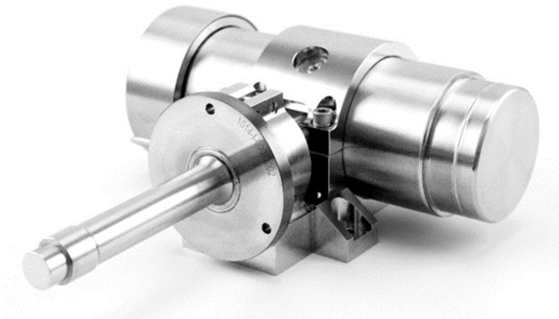

Fig. 5. LPT6510 designed-for-space pulse-tube cooler.

Fig. 6. LPT9310 commercial pulse-tube cooler.

without the need for a moving displacer part.

This means that in a pulse-tube cooler, there is no moving part in the cold finger. Vibration level at the cold side is, therefore, very low and there is no wear-sensitive component left.

By coupling a pulse-tube cold finger with a wear-free compressor, such as a flexure-bearing linear compressor, an extremely high reliability and lifetime can be achieved; as there are no mechanisms left that can cause the cooler to fail, other than the very slow leakage of the working gas (helium). Pulse-tube cryocoolers have been reported to operate over 10 years continuously without any loss in performance.

Pulse-tube coolers are preferred for applications requiring an extended lifetime with a very low failure probability or in applications where maintenance is not possible. Infrared detectors in space are one example. As satellites or spacecraft containing infrared sensors are extremely expensive to launch, they need to operate for an extended period of time with a very high reliability, for which the pulse-tube is ideal.

Other applications for which pulse-tube coolers have found the widespread use include gamma-ray spectroscopy devices which exhibit a high sensitivity to vibrations ("microphonics").

The size, weight, and power (SWaP) consumption of a pulse-tube is typically larger than that of other cryocooler technologies, so this technology is typically used in applications where vibration and reliability are much more important than SWaP.

3.3. Joule-Thomson coolers

The Joule-Thomson (JT) coolers (Fig. 7) are based on the cooling effect generated by free expansion of working gas. The high-pressure gas is brought in a thin tube in the dewar and forced through a calibrated orifice at the end of the tube. Cooled released gas returns to its surroundings exchanging heat through a tube wall with the incoming gas. Incoming gas gets colder and colder over time, until the released liquid reaches a steady state with a fixed stable temperature: the gas boiling point corresponding to the actual gas pressure in the release chamber.

Fig. 7. Joule-Thomson coolers.

There is no need to have a temperature regulation as temperature is fixed by the gas nature. Most common gases are nitrogen (77 K) and argon (88 K), mixture of these gases or air.

Regarded as that, the JT cooler is a total static device with no part moving except the gas.

High-pressure gas supply may be provided from a dedicated compressor or from a finite gas capacity (gas bottle).

If the gas is supplied from a bottle, the duration of the cooling phase will depend on capacity volume and pressure. Adjusting the gas flowrate is then needed to be regulated to the required cooling capacity. Cooling power must be at maximum during the cooling down phase and then can be reduced at steady state. For this purpose, JT coolers may be equipped with a gas flow regulation with a needle moving to partially obstruct the orifice when cold temperature is reached, with a bypass rate being sufficient to compensate thermal loss and to keep steady state. Needle is attached to a gas-filled bellows. When cold temperature is reached, gas inside the bellows turns into liquid and the bellows retraction causes the pin to move into the orifice reducing the gas flow rate.

The gas used for cooling must be ultra-pure as the most common failure mode for a JT cooler is an obstruction in the gas path caused by impurities condensate.

3.4. Thermoelectric coolers (TECs)

This type of cooling is totally static and based on the Peltier effect. When current passes through the junction of two semi-conducting materials, there will be heat absorption or heat rejection depending on current direction.

Thermoelectric modules are built by serial arrangements of semi-conducting bars in a p-type, n-type sequence. After supplying direct current, one side of the module is cooled while the other is heated. Modules can be miniaturized but their efficiency is low, and the biggest problem is with heat dissipation on the hot side as an efficient dissipation mean (fan, water circulation) must be installed on the hot side.

The maximum temperature drop is about 70 °C but colder temperature can be reached by stacking multi stages. Coefficient of performance (COP) is then reduced, and consumption and heat rejection are increased. Cooling components to 200 K is a practical limit for this technology and will require a 4-stage cascade module. While the fundamental work is reported to develop novel thermos-electric materials, in today's technology, the temperature limit is quite fundamental.

4. Using coolers

4.1. Main requirements for cooler selection

While choosing a cooler for an IR detection application, many types of requirements need to be taken into account. Depending on the trade-off to be made between these requirements, the choice will be for one type of cooling.

Main requirements to be evaluated are:

- size and weight,
- consumption, required input power,
- vibration level and acoustic discretion,
- reliability and lifetime,
- cost.

For different types of coolers listed in the above sections, the Table 1 here below proposes an appreciation for each type of requirement. This of course is just a rough appreciation to give some trends and would require a deeper analysis.

Table 1.
Comparing different cooling solutions.

Cooler type	Size & weight	Power	Integration	Vibration and Acoustic	Reliability	Cost
Rotary Stirling	+	+	++	−	0	+
Linear Stirling	0	0	+	+	+	0
Pulse-tube	−	−	0	++	++	−
Joule-Thomson	+	++	−−	+	+	+
Thermoelectric	++	−−	−−	++	++	+

Rotary Stirling coolers can be miniaturized with a low mass (down to 150 g). They are easy to integrate, as they usually do not require a specific thermal exchanger.

Linear Stirling coolers show low level of induced vibration and long lifetime, especially for versions including flexure bearings.

Pulse-tube coolers show superior behaviour in both induced vibration and reliability aspects. There are no moving parts in the cold finger.

JT coolers have few moving parts (none for unregulated versions), reliability is improved. A high-pressure gas source is required.

TECs can be miniaturized and are cheaper than any other system though thermal dissipation means must be included in the global price. They cannot be used for cooling under 200 K.

4.2. Importance of dewar configuration

In the early days of IR detection, Stirling coolers were delivered complete with a cold finger, helium-filled for life. These coolers were slipped to be assembled in the dewar cold finger. This type of assembly is known as the "slip-on" configuration.

Then, it appeared that this configuration was not optimum for heat transfer and another alternate configuration has emerged: the integrated dewar cooler assembly (IDCA).

To reduce heat losses, the cold finger becomes a common part between dewar and cooler which significantly reduces heat losses due to thermal conduction in the cold finger wall, as there is only one wall now instead of two. Moreover, having one shared cold finger instead of two also eliminates the gap between the two cold fingers in the slip-on configuration. This gap necessary for mechanical purpose generates thermal resistance. Due to this thermal resistance, the cooling temperature at cooler boundaries must be lower than the set point to compensate the temperature difference induced by the interface resistance (Fig. 8).

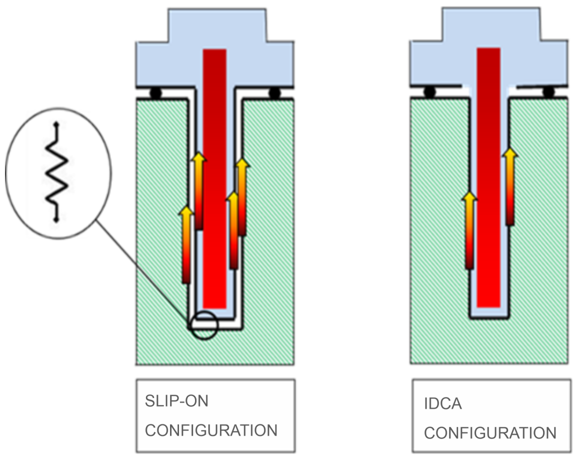

Fig. 8. Slip-on and IDCA configuration.

The cooler is delivered incomplete with a protector covering displacer (no cold finger). Cooler integration is then done in the dewar with the dewar cold finger covering cooler displacer to complete the cooler outer envelope. Final gas filling is done at this stage.

This IDCA configuration has brought much improvement by drastically reducing heat losses and consequently heat load.

After introduction of the IDCA configuration, significant improvements have been achieved in size, weight, and consumption of the coolers giving way to the miniaturized series of coolers that emerged in the 1990s and allowing the use of cryocoolers in handheld thermal imagers (HHTI) [2]. It should be noted that the IDCA configuration requires the cooler to be open to air during integration. To prevent contamination and internal corrosion in this phase, this integration must be done in a controlled environment under strict precautions to avoid contamination.

4.3. Cooling sequence

This section describes a typical cooling pattern of a rotary Stirling cryocooler. Same pattern will more or less apply to other cooling means starting with a full cooling power stage for cooling down and onset of a regulation phase approaching temperature set point value.

The electronic driver is switched to the DC power supply. A cold plane temperature sensor (diode) is connected to the driver. According to Fig. 9 at power on (1), the cooler starts to run at full speed and temperature instantly drops. Full speed regime is kept until reaching the temperature set point. Input power slightly increases as temperature gets colder. When cold temperature set point is reached (2), the regulation phase is launched, and the driver modulates voltage supply [pulse width modulation (PWM)] to reach and keep the constant cold temperature at the set point value. Rapidly, temperature, speed, and consumption reach stable values and remain stable in time (3). Smoothness of temperature control and temperature hysteresis depend on a PID setting of the driver that can be adapted to an application parameter: heat load, thermal mass (Fig. 9). Temperature can be kept stable to ±0.2 K this way.

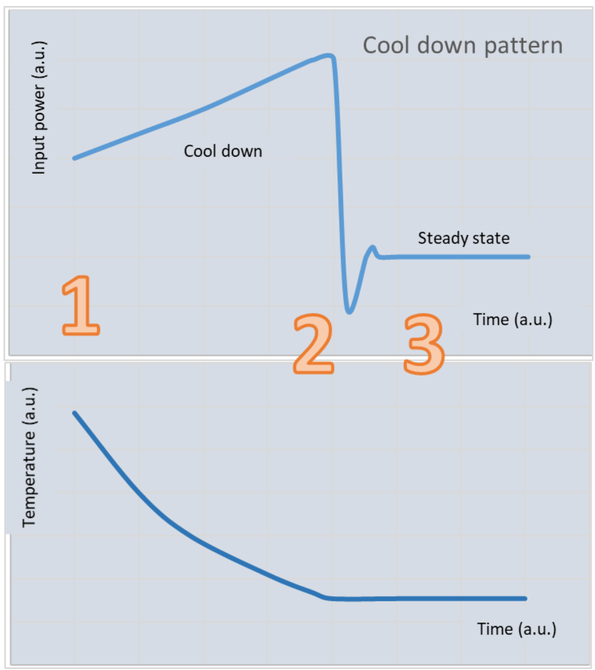

Fig. 9. Typical cooling pattern.

4.4. Lifetime and reliability

TEC, JT, and pulse-tube coolers are definitely highly reliable and can be recommended for critical and intense uses.

Linear Stirling coolers show a very good lifetime especially if they are built using flexure bearings. In this case, compressor pistons are guided by the flexure bearing and there is no mechanical contact, hence, no friction between piston and cylinder. Mechanical wear out is much reduced.

Rotary Stirling coolers appeared in the past as low-cost solutions with reduced lifetime. Many improvements have been made recently and Thales have released results from the lifetime test reporting MTTF values over 50 000 hours for RM2 coolers. Choice of rotary coolers for 24/7 applications can now be considered [3].

4.5. Drive electronics

An electronic drive has to be connected to the rotary Stirling cooler to command voltage supply distribution according to the relative rotor/stator position as measured by the Hall effect sensor signal. Voltage is switched from one stator winding to the next every time a sensor detects a magnetic polarity inversion in the rotor rotation.

The driver also manages the voltage supply to regulate the cold temperature to the required set point. Voltage modulation is performed through PWM, the rotary speed is a direct function of the supply voltage. Cooler rotary speed represents the number of thermodynamic cycles performed per unit of time, hence the cooling power.

Historically, drive electronics for Stirling coolers were analogue, but as with most other electronics, digital circuits have brought many advances.

The SWaP from analogue to digital cooler drive electronics improved size and efficiency with much better stability for temperature in steady state regulation through the entire ambient temperature range. Communication can be set with a computer, and this is quite convenient for the set point input and the cooler monitoring as main cooler parameters: voltage, current, speed, temperature can be monitored and recorded in real time.
Hour counter can be included. As the power supply of a cryocooler might be a source of electromagnetic interference (EMI) disturbing detector signal reception, filters can be commonly placed on the driver circuit board to tackle these issues.

5. RMs1 SWaP cooler

The last rotary cooler developed by Thales is the RMs1 (Fig. 10). This cooler is dedicated to HOT detectors. It is now qualified and in serial production. The conclusion of this qualification effort is that the RMs1 is, due to its high efficiency and power density, able to cool a very wide range of applications from at least 110 K to 170 K. The RMs1 product has been designed to deliver a high cryogenic power at a high overall efficiency within half the volume and half the weight of the RM2 cooler. Moreover, the RMs1 implemented the generic cold-finger interface designed to reduce the variation in dewars at our customers.

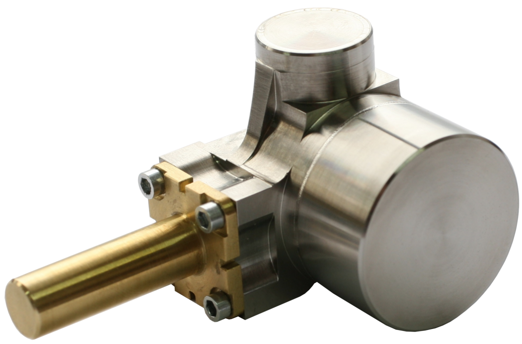

Fig. 10. RMs1 rotary cooler.

This paragraph briefly summaries the characteristics and the performances of the RMs1. This cooler is a very small cooler designed for HOT applications (such as 150 K).

Size+compact cooler: 72×42×42 mm

Weight+light cooler: 142 g

Power+efficient cooler: 1 W @150 K

High efficiency 110 =>170 K/0.8 W @150 K

Quiet cooler: noise: silent @15 m
Induced vibrations < 40 mNrms

Reliable cooler: 15 000 hr. @63% of failure

Short cool down time: < 2 min. @150 K

The conclusion of this chapter is that the RMs1 is a very stable product coupling, compact, high efficiency with low-induced vibrations and acoustic noise levels. This is the best solution where SWaP is the key criterion. Due to its high achievable cooling power, the RMs1 is also a good platform for many other applications with larger detectors and detectors operating at lower temperature (Fig. 11). This allows to use a single type of cooler in a large variety of applications, thereby optimizing storage, logistics and manufacturing costs [4].

Fig. 11. NEO electronic driver.

It can be noticed that a new electronic driver was developed by Thales in order to reduce the SWaP consumption of the cooler and electronic assembly.

6. Application for RMs1 cooler @ IRnova

The RMs1 cooler is an excellent choice for detectors for which low SWaP consumption is required. Applications which pull demand for volume production of such detectors are portable, handheld, battery-powered or more simply space-constrained applications in both industrial and surveillance/defence applications.

The first European type II superlattice (T2SL)-based low SWaP IDDCAs (Oden MW, Fig. 12) includes the RMs1 as its cooling unit [5]. In the Oden MW detectors, the RMs1 coolers are combined with dewars with F/4 and F/5.5 optical configurations and with VGA format detectors (640×512 pixels, 15 μm pitch). This enables a very compact design of the IDDCA with a weight of 230 g and size of 48×44×98 mm. The Oden MW IDDCA is typically operated in the temperature range from 120 K to 150 K with a power consumption decreasing from 3 W at 120 K to values below 2 W when further increasing the operating temperature (Fig. 13). The cooling down time also decreases as the operating temperature is increased, from 3 min. at 120 K to 2.5 min. at 150 K (Fig. 14).

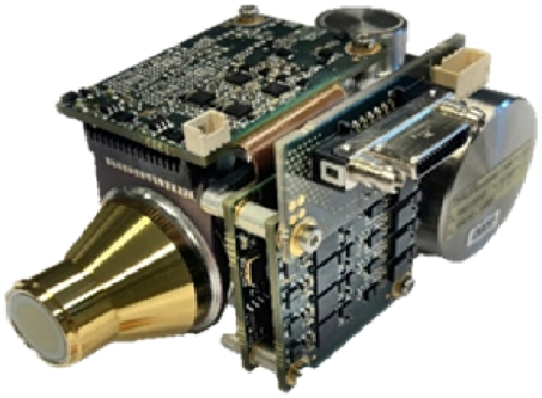

Fig. 12. Oden MW SWaP IDDCA with RMs1 cooler in production at IRnova.

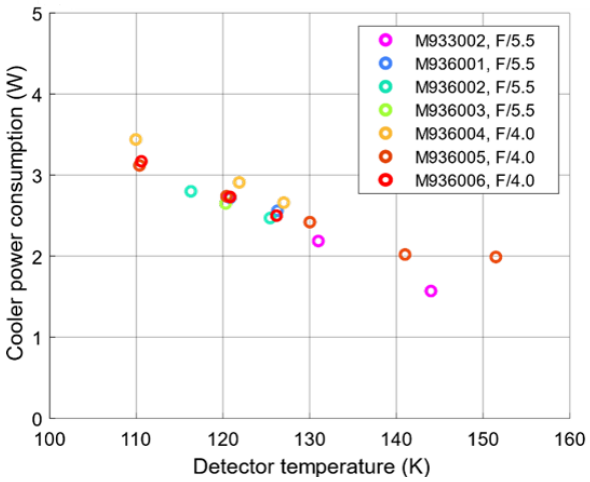

Fig. 13. Steady-state power consumption of seven different Oden MW IDDCAs, in the temperature range from 110 K to 150 K.

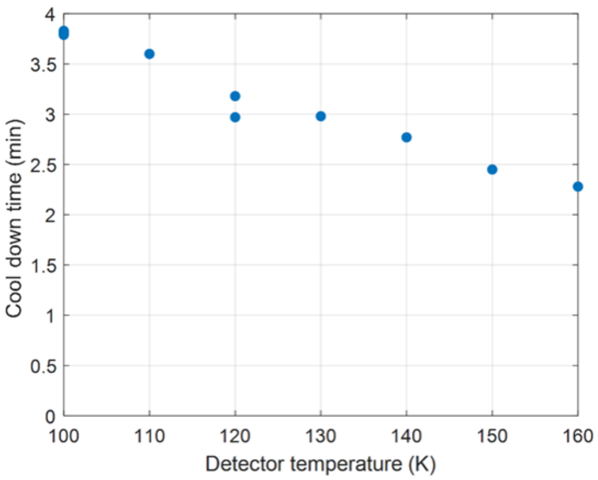

Fig. 14. Cool down time of an Oden MW IDDCA in the temperature range from 100 K to 160 K.

7. Conclusions

Infrared detector cooling to cryogenic temperatures remains a necessity for demanding IR detection applications. Progress in IR detector technologies now allows operation at a higher temperature up to 150 K instead of the historical 77 K. New coolers dedicated to these HOT applications are smaller, lighter, with a lower power consumption.

References

[1] Rogalski, A. History of infrared detectors. *Opto-Electron. Rev.* **20**, 279–308 (2001). https://doi.org/10.2478/s11772-012-0037-7

[2] Vasse, C., Griot, R., Abousleiman, V. & Benschop, T. SWaP approach on Thales rotary cryocoolers leading to environmental impact improvements. *Proc. SPIE* **12107**, 121070J (2022). https://doi.org/10.1117/12.2618790

[3] Cauquil, J. M., Seguineau, C., Vasse, C., Raynal, G. & Benschop, T. Lifetime validation of high-reliability (>30,000 hr) rotary cryocoolers for specific customer profiles. *Proc. SPIE* **10626**, 106260H (2018). https://doi.org/10.1117/12.2309792

[4] Vasse, C., Martin, J. Y., Cauquil, J. M., Raynal, G. & Benschop, T. Optimized tuning of RMs1 for wide range of applications. *Proc. SPIE* **11002**, 110020A (2019). https://doi.org/10.1117/12.2520382

[5] Delmas, M. *et al*. HOT SWaP and HD detectors based on Type-II superlattices at IRnova. *Proc. SPIE* **12107,** 121070R (2022). https://doi.org/10.1117/12.2618752