MEDICAL PHANTOM DISCRIMINATION WITH HYPERSPECTRAL IMAGING

Carolina Blanch-Perez-del-Notario[a], Ferran Soebrata[b], Glenn Beukema[b], Steven Thijs[a] and Murali Jayapala[a]

[a]Imec, Kapeldreef 75, 3001, Leuven, Belgium; [b]Quest Photonic Devices B.V., Quest Medical Imaging B.V. Westrak 3, 1771 SR Wieringerwerf, The Netherlands.

ABSTRACT

Shortwave infrared (SWIR) imaging offers deeper tissue penetration, reduced scattering, and sensitivity to absorbers such as water and lipids, making it promising for fluorescence-guided surgery. Yet clinical use remains limited by system cost, complexity, and the lack of fast, application-specific solutions. In the RETINA European Project, we assess real-time multispectral snapshot SWIR imaging for surgical guidance using tissue-mimicking phantoms. High-resolution SWIR hyperspectral imaging identifies the most discriminative wavelengths for phantom separation, enabling selection of a reduced band set. We then emulate broader quantum-dot-like spectral responses and complement this with preliminary measurements from quantum-dot SWIR sensors centered at 1200 and 1450 nm. Results indicate that both compact multispectral narrow-band InGaAs and broad-band quantum-dot-based cameras can retain useful discrimination contrast, supporting snapshot SWIR sensing as a path toward higher-resolution, lower-cost, application-tailored systems for real-time surgery.

***Index Terms*—** Hyperspectral, SWIR, medical phantom, snapshot imaging, quantum-dot.

1. INTRODUCTION

Shortwave infrared (SWIR, ~1000–2000 nm) imaging and spectroscopy are gaining momentum in biomedicine because they can provide complementary contrast mechanisms to visible/near infrared (NIR) approaches, driven by reduced scattering and increased sensitivity to endogenous absorbers such as water and lipids [1],[2]. The authors in [2] survey rapidly expanding application areas—including dental imaging, otology, and tumor-margin assessment—while also emphasizing practical barriers that have historically slowed clinical translation, notably detector cost/availability, illumination constraints, and system complexity ([1], [2]). In parallel, methodological advances are accelerating interpretation of high-dimensional SWIR data: in [3] a review of machine learning-assisted SWIR techniques is presented and the role of data-driven models for robust tissue characterization and, ultimately, personalized decision support is highlighted. For fluorescence-guided surgery, SWIR detection of SWIR wavelength emission tails of established NIR dyes has been proposed as a pragmatic route to translate SWIR advantages without requiring entirely new contrast agents [1]. In an in vivo neuroblastoma setting, the report in [4] presents an improved tumor visualization and margin delineation using a multispectral SWIR setup but also identify long exposure times as a key obstacle for real-time intraoperative use. In contrast, in [5] a systematic NIR-versus-SWIR comparison on ex vivo clinical tumor samples labelled with cetuximab-IRDye800CW is performed, showing that SWIR does not yet deliver a consistent advantage over NIR, with tissue-dependent background signals limiting performance—although an increase in tumor-border contrast is observed in specific cases. These studies indicate that translation depends not only on wavelength choice but also on practical system design (speed, noise, illumination, and calibration) and on standardized validation protocols.

A recurring need across SWIR modalities is reliable, application-relevant validation using tissue-mimicking phantoms that decouple optical effects from biological variability and enable repeatable benchmarking of acquisition strategies and algorithms. The work in [6], for example, demonstrates phantom-based validation of a compact SWIR diffuse-reflectance probe for hydration-sensitive measurements in gelatine–Intralipid constructs with controlled scattering and water content, illustrating how carefully engineered phantoms can support quantitative performance claims. However, much of the SWIR literature—especially in imaging and point spectroscopy surveys—still relies on “line-scan” or otherwise slow acquisition schemes [2], which constrains deployment in dynamic and intraoperative scenarios. The key gap we target in the RETINA European Project [7] is the limited use of multispectral snapshot SWIR cameras for clinically realistic, real-time surgical guidance. In this work, we first use recipe-controlled medical phantoms for initial validation and repeatable performance assessment, and we leverage high-resolution SWIR hyperspectral imaging to identify the most discriminative wavelengths for tissue (or target) separation. These selected bands then provide a concrete pathway toward customized, application-specific sensor designs—bridging

the gap between promising SWIR contrast mechanisms and practical, real-time imaging systems for surgery. In this respect, Quantum-Dot (QD) based cameras [8-10] are emerging as high spatial resolution snapshot spectral sensors in the SWIR range with potentially lower manufacturing cost. Because silicon is largely transparent in the SWIR range, conventional SWIR cameras typically use costly III–V detectors such as InGaAs, whose pixel pitches of about 5 µm or more limit spatial resolution. Although still at an early technology-readiness stage, QD technology offers a promising lower-cost alternative, with demonstrated pixel pitches below 2 µm and compatibility with monolithic CMOS-wafer processing, enabling higher pixel densities and potentially lower manufacturing costs [8–10]. Within the RETINA project, new QD detectors are being integrated into and evaluated with a modular SWIR camera prototype supporting video acquisition at up to 30 fps. While current QD sensors have a single peak spectral response centered around a selected wavelength of interest, future sensor developments are targeting a few spectral bands on the same sensor [10]. In our work, we use a high-resolution SWIR hyperspectral camera with 100 bands to first identify the most relevant spectral bands for our phantom discrimination. We then mimic a broader QD-alike band response from our high-resolution data and perform some validations with actual QD camera and sensors centered around specific wavelengths. To our knowledge, this is one of the first works exploring the feasibility of phantom discrimination in the SWIR range with multispectral snapshot SWIR cameras and with Quantum-Dot cameras.

2. MATERIALS AND METHODS

We use a Snapscan camera [11] in the short-wave infrared (SWIR) range offering 100 bands of 5-nm width in the 1100 and 1650 nm range and a maximum spatial resolution of 1200 x 640 pixels. Thanks to its internal translation stage, there is no need for an external scanning movement. Scanning is handled internally, performing data-set acquisition as easily as with a snapshot camera. We also use the imec SWIR Snapshot camera [12]. This camera offers the benefits of mosaic snapshot cameras at the expense of a reduced spatial-spectral resolution with respect to the Snapscan SWIR camera. The mosaic sensor is based on a 3x3 mosaic filter patterning on a SWIR detector, implemented by means of high-transmission Fabry-Pérot interferometers processed using thin-film technology [11]. This SWIR camera offers 9 bands in the 1100-1700 nm range at a spatial resolution of 213 x 170 pixels, which can be up-sampled to 633x504 pixels. Both cameras were placed approximately 30 cm above the samples and were used in combination with a Navitar 25 mm lens and an illumination setup consisting of 4 halogen Osram lamps (12v 2800k 20W).

To evaluate the suitability of SWIR snapshot cameras for real-time medical applications, such as tumor resection, tissue-mimicking phantoms were developed to reproduce relevant body-tissue properties in the SWIR range. Additives were used to determine whether they could be distinguished in the SWIR range. Such phantoms are widely used in medical imaging to assess device performance under controlled and reproducible conditions. In this study, the phantoms consisted of a base material [13] combined with different concentrations of commonly used scattering, absorbing, or fluorescent additives, following [14]. This design enabled an individual assessment of the optical contribution of each component in the SWIR range, for which limited information is currently available. Phantoms were preferred over real tissue to improve reproducibility and to avoid unnecessary burden or risk to patients during feasibility testing. The phantom design was based on the approach reported in [14], which uses 3D printing by selective UV curing of a resin. In that work, scattering, absorbing, and fluorescent agents are embedded in a transparent resin to create materials with tunable optical properties [15]. Eight phantoms (A1–A8) were produced in three geometries: small, elongated cuboids (30x10x20 mm), wider cuboids (30x30x20mm), and textured phantoms (30 mm diameter) intended to better approximate the shape of removed biological tissue. The labels A1 to A8 indicate different additive concentrations in the base material. Table 1 summarizes the additive concentrations used in the phantom formulations, and examples of the phantoms employed for medical validation are shown in Figure 1.

Table 1: : Additive concentrations in the phantom base materials, as mass of additive per mass of base resin.

Phantom	TiO_2 (scatterer)	Hemin (absorber)	Nigrosin (absorber)	ICG (fluorophore
A1	0	0	0	0
A2	1.4 mg/g	0	0	0
A3	6.5 mg/g	0	0	0
A4	0	40 µg/g	0	0
A5	0	0	0.7 µg/g	0
A6	0	0	3.5 µg/g	0
A7	0	0	0	23 µg/g
A8	0	0	0	0.69 µg/g

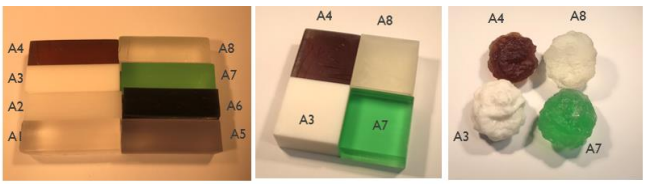


Figure 1: Resin phantoms: small cuboids (left), bigger wider cuboids (middle) and textured phantoms (right).

We first define a ground truth by creating masks selecting different phantom material types. From the masks created 2000 spectra samples per material class are randomly selected. Out of these 2000 spectra per material, 20% are randomly selected for training and the remaining 80% for validation. We use PerClass software [16] in combination with MATLAB processing software [17] and use Linear Discriminant Analysis (LDA) [18], as a feature selection step, followed by a Quadratic Discriminant Classifier (QDC) [18]. This is a pixel-based classifier; therefore, we do not use any spatial information to support classification. To find the most discriminative bands in our wavelength range we use an in-house developed Genetic Algorithm [19],[20] in combination with the classifier. The algorithm aim is to find, based on the reflectance spectra of each band, a near-optimal subset of bands providing the highest mean classification accuracy in the test set. This can steer the development and wavelength selection of next-generation imaging sensors, such as emerging Quantum Dot (QD) devices. At present, these sensors are generally broadband or tuned to a single peak wavelength [8], but future developments are expected to enable the integration of spectral filters targeting several selected wavelength bands [10]. This would open the door to low-cost, high-spatial-resolution multispectral QD sensors and could accelerate broader adoption of this technology. For this reason, identifying the most relevant bands for each application is an important step toward the effective design and use of future QD sensors. The relevant-band analysis in this study is based on high-resolution Snapscan SWIR data spanning 1100–1680 nm with more than 100 narrow bands (about 5 nm full-width-half-max (FWHM)) [21]. The selected bands are intended to guide the design of QD sensors for medical applications. Because QD sensors typically have a much broader spectral response, around 100 nm FWHM, as seen in Figure 2 we also assess whether bands that are highly discriminative in narrowband hyperspectral data remain effective when implemented as broader responses that may blur fine spectral features.

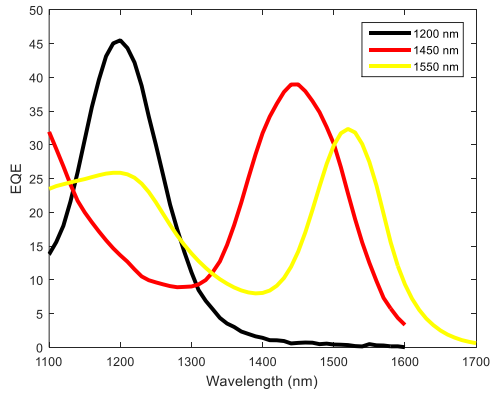


Figure 2: Typical EQE (external quantum efficiency) of QD sensors

For this reason, we approximate realistic broad QD spectral responses by Gaussian filters and study as well simulated bandwidths of 20, 50, and 100 nm FWHM to explore different performance tradeoffs. The comparison is made for the single most relevant band as well as for the best two- and three-band combinations.

3. RESULTS AND DISCUSSION

Figure 3 shows the mean SWIR phantom reflectance spectra, including a resin-related peak around 1300 nm. Using the full 100-band Snapscan data, the phantoms are classified with high accuracy: 92.75% mean pixel accuracy and 87.42% minimum accuracy for A5.

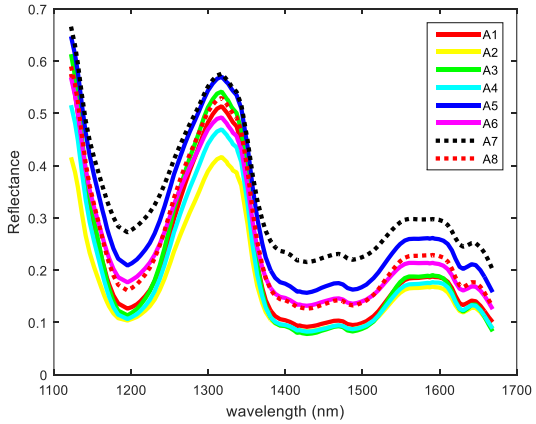


Figure 3: Mean reflectance spectra of phantom tissues.

Our genetic-algorithm band-relevance analysis identified then the most discriminative SWIR bands, summarized in Table 2. Simulating the 9-band Snapshot SWIR sensor from the Snapscan data retained high performance, reaching 91.1% mean accuracy and 77.3% minimum accuracy (A5 phantom). To validate this, we performed measurements with the Snapshot SWIR camera, which confirmed the expected high-accuracy of the 9-band configuration. Moreover, the results in Table 2 suggest that fewer optimized bands, down to four, can still provide useful discrimination with mean pixel accuracies of 82% and minimum of 62%

Table 2: Band relevance analysis for Snapscan SWIR

BandSubset	Mean/Min(%)	Bands (nm)
All 101	92.7%/87.4%	1122 … 1668
Best 9 bands	94.9%/86.4%	1131,1141,1344,1363,1460, 1493,1605,1636,1663
Snapshot 9 bands	91.1%/77.3%	1122,1131,1152,1183,1201,1288 1374,1452,1641
Best 6 band	91.8%/76.6%	1122,1146,1317,1357,1548,1663
Best 5 band	86.2%/61.6%	1127,1141,1336,1440,1544
Best 4 band	81.0%/62.0%	1131,1351,1370,1480
Best 3 band	71.9%/30.5%	1127,1351,1593
Best 2 band	61.1%/24.9%	1141,1668

Figure 4 shows the classified image based on the best four bands identified, revealing good discrimination, with most errors occurring near phantom borders.

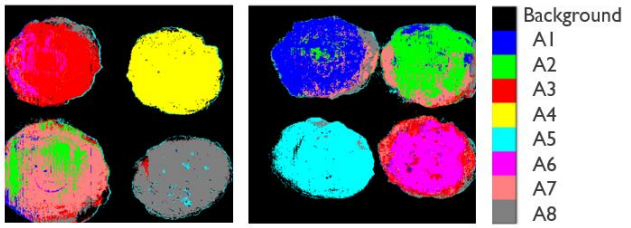


Figure 4: Classified images of best 4 narrow bands.

Next, we used the 100-band, high-spectral-resolution dataset, with bands spaced at 5 nm intervals, to model broad QD-like spectral responses. Gaussian filters with full widths at half maximum (FWHMs) of 20, 50, and 100 nm were applied to the most relevant individual band and the best two-band combinations. For the 1122 nm band, broader responses could not be fully simulated because the measured spectral range does not extend below 1122 nm. The resulting classification performance is reported in Table 3. We also simulated the performance of currently available QD sensors centered at 1200 and 1420 nm, together with a potential 1600 nm sensor that is manufacturable but has not yet been produced. This analysis focuses on the flat, cube-shaped phantoms because they allow more straightforward validation with later measurements from QD sensors.

Table 3: Impact of Full-Width-Half-Max of different available quantum dot spectral responses

QD (nm)	Mean/min @100nm	Mean/min @50nm	Mean/min @20nm	Mean/min 5nm
1668	66.0/22%	63.6%/10%	58.5%/11%	**68.0%**/30%
1200	**64.8**/12%	62.7%/20%	54.5%/29%	55.3%/ 2%
1420	60.1/ 17%	**64.0%**/19%	61.4%/20%	59.8%/ 24%
1600	65.1%/ 23%	**66.0%**/14%	63.4%/10%	63.5%/ 12%
1122 1162	94.0%/ 79%	93.3%/ 78%	93.2%/78%	**94.2%**/ 77%
1200 1420	81.8%/ 57%	**86.8**%/ 60%	74.5%/ 35%	71.8%/ 34%
1420 1600	**88.2%**/58%	81.4%/ 42%	82.7%/ 42%	82.6%/ 45%
1200 1600	**88.5%**/58%	82.7%/ 48%	73.3%/ 27%	74.1%/ 35%

As expected, the wavelengths identified by the algorithm as optimal for discrimination (1668, 1122, and 1162 nm) achieved the highest accuracy when modelled with narrow responses of 5 or 20 nm, which preserve spectral specificity. By contrast, our currently available QD sensor wavelengths of 1200, 1420, and 1600 nm benefited from broader responses of 50–100 nm, likely because they are not aligned with the optimal bands. For these wavelengths, a 50 nm FWHM provided the best overall trade-off. A 20 nm FWHM offered limited benefit and could reduce accuracy when the QD response peak was not centered on an optimal band. As shown in Figure 5, the simulated 1200 and 1420 nm QD sensor bands with a 50 nm FWHM nevertheless achieved reasonable discrimination. Segmentation and object level-based approaches could further increase accuracy in this case.

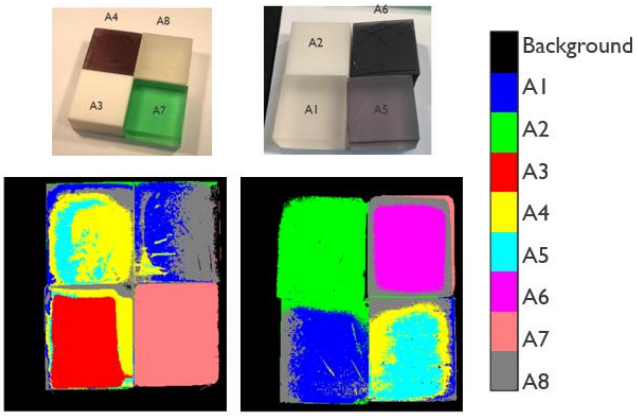


Figure 5: Classified image for simulated 1200/1420 QDs

To validate this assumption, these cuboid phantoms were measured independently with two QD sensors centred at 1200 and 1420 nm, each with approximately 100 nm FWHM. Simulation in Table 3 suggests that these two broad bands could provide sufficient discrimination capability. Figure 6 shows the phantom images acquired with the 1200 nm sensor (left) and the 1420 nm sensor (right).

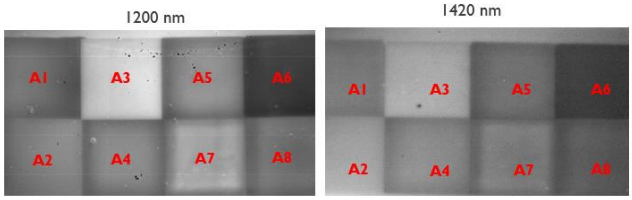


Figure 6: QD images of phantoms at 1200 and 1420 nm.

The single broadband response at 1200 nm allows partial discrimination of phantoms A3, A5, A6 and A7, while A1, A2, A4 and A8 overlap more strongly. Similarly, the 1420 nm sensor can partly discriminate between A2, A3 and A6, but several phantoms remain overlapping and prone to misclassification. We then merged the 1200 and 1420 nm QD images to emulate a future two-band QD sensor. This resulted in a mean accuracy of 65.5%, with A8 mainly misclassified as A3 or A5 and A4 mainly misclassified as A1. To our knowledge, there is no accepted regulatory benchmark for phantom discrimination accuracy using spectral imaging. Regulatory evaluation is generally performed at the level of the final clinical intended use. Thus, while two broad bands provide some discrimination, 3–4 QD bands are likely needed to reach higher accuracy.

4. CONCLUSIONS

Our study on spectral band-relevance analysis confirmed that accurate medical phantom discrimination (>90%) can be achieved with only a few spectral SWIR bands, supporting the use of snapshot multispectral imaging in the SWIR range. This was validated with Snapshot SWIR measurements for narrowband imaging and further investigated through simulations and measurements with broadband QD sensors. Overall, the results indicate that both narrowband snapshot cameras and broadband QD sensors are promising for compact spectral imaging systems in medical domain providing a few selected bands. Future work will focus on manufacturing improved phantoms and developing additional QD sensors at the relevant wavelengths that further increase discrimination accuracy.

ACKNOWLEDGEMENTS

We want to acknowledge the Retina project, European Union's Horizon Europe research and innovation programme under grant agreement No 101135529, for which this research has been done.

REFERENCES

[1] Carr, J. A. (2018). Shortwave Infrared Imaging and Its Translation to Clinically Relevant Designs (PhD thesis).

[2] Roblyer, D., Pilvar, A., Pham, T., Spink, S., … & Pierce, M. C. (2025). Review of shortwave infrared imaging and spectroscopy in tissue. Biomedical Optics Express, 16(12), 5028-5062.

[3] Salimi, M., Roshanfar, M., Tabatabaei, N., & Mosadegh, B. (2023). Machine learning-assisted short-wave infrared (SWIR) techniques for biomedical applications: towards personalized medicine. Journal of Personalized Medicine, 14(1), 33.

[4] Privitera, L., Waterhouse, D. J., ... & Giuliani, S. (2023). Shortwave infrared imaging enables high-contrast fluorescence-guided surgery in neuroblastoma. Cancer Research, 83(12), 2077-2089.

[5] Keizers, B., Nijboer, T. S., ... & van der Zaag, P. J. (2025). Systematic comparison of fluorescence imaging in the near-infrared and shortwave-infrared spectral range using clinical tumor samples containing cetuximab-IRDye800CW. Journal of Biomedical Optics, 30(S1), S13708.

[6] Farouq, G., Vyas, D., & Tofghi Zavareh, A. (2026). Development and phantom validation of a small-form-factor SWIR emitter probe for hydration-sensitive spatial-ratio measurements in gelatin-Intralipid phantoms. Sensors.

[7] Photonic integrated LIDAR and snapshot spectral imagers for scaling up multimodal perception in precision applications | RETINA | Project | Fact Sheet | HORIZON | CORDIS | European Commission. https://cordis.europa.eu/project/id/101135529 (accessed May 29, 2026).

[8] V. Pejović et al., "Infrared Colloidal Quantum Dot Image Sensors," in IEEE Transactions on Electron Devices, vol. 69, no. 6, pp. 2840-2850, June 2022, doi: 10.1109/TED.2021.3133191

[9] P.E. Malinowski, V. Pejović.,… & J. Lee, "Colloidal quantum dot image sensors: a new vision for infrared", 2022 IEEE International Electron Devices Meeting, paper 19.3 (2022).

[10] Kara, G., Puybaret, R., Malainou, A., ... & Pejovic, V. (2025, December). Metasurface Cointegration with Colloidal Quantum Dot Photodiodes Enabling Facile Spectral Sensors. In 2025 IEEE International Electron Devices Meeting (IEDM) (pp. 1-4). IEEE.

[11] Gonzalez, P., Pichette, J., ..., & Lambrechts, A. (2018, May). "An extremely compact and high-speed line-scan hyperspectral imager covering the SWIR range". In Image Sensing Technologies: Materials, Devices, Systems, and Applications V (Vol. 10656).SPIE. https://doi.org/10.1117/12.2304918

[12] Blanch-Perez-del-Notario, C., Luthman, S., Lefrant, R., Gonzalez, P., & Lambrechts, A. (2022, May). Compact high-speed snapshot hyperspectral imager in the SWIR range (1.1-1.65 nm) and its potential in sorting/recycling industry. In Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXVIII (Vol. 12094, pp. 47-55). SPIE.

[13] Elegoo ABS-like resin transparent. Shenzhen Yongchanghe Technology Co., Ltd., Shenzen, China. https://eu.elegoo.com.

[14] Ruiz, A. J., Garg, S., … & Pogue, B. W. (2021). 3D printing fluorescent material with tunable optical properties. Scientific Reports, 11, 17135. https://doi.org/10.1038/s41598-021-96496-0

[15] LaRochelle, E. P. M., Streeter, S. S., … & Ruiz, A. J. (2023). 3D-printed tumor phantoms for assessment of in vivo fluorescence imaging analysis methods. Molecular Imaging and Biology, 25, 212–220. 10.1007/s11307-022-01783-5

[16] PerClass BV 2008-2022, Delft, NL. http://perclass.com/perclass-toolbox/product

[17] Matlab (2019). The Mathworks, Natick. https://mathworks.com.

[18] Naes, T., Isaksson, T., Fearn, T. & Davies, T. "A User-Friendly Guide to Multivariate Calibration and Classification". NIR Publications (2004).

[19] Holland, J. H. (1992). Genetic algorithms. Scientific American, 267. https://doi.org/10. 1038/scientificamerican0792-66.

[20] Blanch-Perez-del-Notario, C., Baert, R., & D'Hondt, M. (2012). Multi-objective genetic algorithm for task assignment on heterogeneous nodes. Int Journal of Digital Multimedia Broadcasting, 2012, 1-12.

[21] Full width at half maximum - Wikipedia