



HAL
open science

From compressed sensing to learned sensing with metasurface imagers

Philipp del Hougne

► **To cite this version:**

Philipp del Hougne. From compressed sensing to learned sensing with metasurface imagers. Passive and Active Millimeter-Wave Imaging XXIV, Apr 2021, Orlando, France. pp.11, 10.1117/12.2587774 . hal-03849879

HAL Id: hal-03849879

<https://hal.science/hal-03849879>

Submitted on 12 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

From compressed sensing to learned sensing with metasurface imagers

Philipp del Hougne

Univ Rennes, CNRS, IETR - UMR 6164, F-35000, Rennes, France

ABSTRACT

Reducing the latency of electromagnetic imaging is a crucial objective for applications such as security screening, autonomous driving, and touchless human-machine interaction. In that respect, a fundamental caveat with conventional compressed sensing techniques is that initially all information is indiscriminately multiplexed across a diverse set of measurement modes, and only during data processing one begins to select the information that is relevant to the task (e.g. concealed weapon detection). In order to only acquire relevant information in the first place, and hence drastically reduce the number of necessary measurements, the “learned sensing” paradigm suggests to interpret reconfigurable measurement hardware (e.g. a dynamic metasurface aperture) as a trainable physical layer. The latter can be directly integrated into the machine-learning pipeline used on the data processing side such that one can jointly optimize physical weights (measurement settings) and digital weights (processing network). We discuss our recent numerical and experimental investigations of this new approach to electromagnetic imaging. Our results show that a considerable reduction of the number of scene illuminations is possible by using learned illumination patterns as opposed to conventional patterns (random, orthogonal or principal scene components). At the same time, we find that there are no intuitive explanations for the learned patterns. We also clarify whether the resolution of sub-wavelength features in the scene is limited by the conventional diffraction limit.

Keywords: Intelligent Imaging, Learned Sensing, Programmable Metasurface Imager, Latency, Compressed Sensing, Machine Learning

1. INTRODUCTION

Latency currently limits the most promising applications of electromagnetic imaging and sensing, including security screening, autonomous driving and touchless human-machine interactions. Every imaging pipeline has two main potential latency bottlenecks: (i) the required number of measurements per imaged scene, and (ii) the processing burden. Many efforts to reduce latency revolve around the idea of “compressing” sparse scene information to reduce the number of measurements.¹ An underappreciated point of view regarding latency is that of the flow of information: at what point in the pipeline do we separate relevant and irrelevant information? The vast majority of current electromagnetic imagers indiscriminately acquires all information, including large amounts of irrelevant information, and then attempts to identify relevant pieces of information during processing. The double disadvantage of this strategy is immediately obvious: more measurements are needed, and it takes more processing effort to select relevant information. Here, we outline a generic strategy for identifying relevant information already during the measurement process, and demonstrate its significant latency gains in various prototypical case studies.

To develop our ideas, without loss of generality we focus in the following on two imaging architectures based on programmable metasurfaces. Both can be thought of as low-cost alternatives to the use of coherently controlled antenna arrays for pattern synthesis. While the latter require countless radiofrequency chains (one per antenna), both metasurface paradigms that we consider operate with a single feed. In one instance, a horn antenna illuminates a programmable metasurface acting as reconfigurable reflect-array that impresses a desired pattern on the reflected wavefront. In another instance, the single feed excites a two-dimensional waveguide which is patterned with metamaterial elements that are programmed to leak energy at desired locations from

Correspondance to P.d.H. (philipp.del-hougne@univ-rennes1.fr).

the waveguide to the far field in order to holographically synthesize a desired pattern. While the second instance is more compact, both have in common the use of *in situ* programmable meta-atoms to shape the waves that illuminate the scene. From an imaging perspective, both constitute programmable coded apertures suitable for multiplexed measurements with configurational diversity.

2. INFORMATION FLOW

To start our analysis of the information flow in electromagnetic imagers, we note that in the vast majority of cases, creating a visual image of the scene is not the ultimate goal but an intermediate step. The image is then further analyzed (typically with artificial intelligence tools) to detect concealed weapons (security screening), to identify vehicles or pedestrians on the road ahead (autonomous driving), or to recognize hand gestures (touchless human-machine interaction). Often these ultimate processing goals are neglected and studies content themselves with presenting an uninterpreted visual image of a scene; yet it is the specific interpretation task that determines which information is relevant, and which information is acquired but useless for the task at hand.

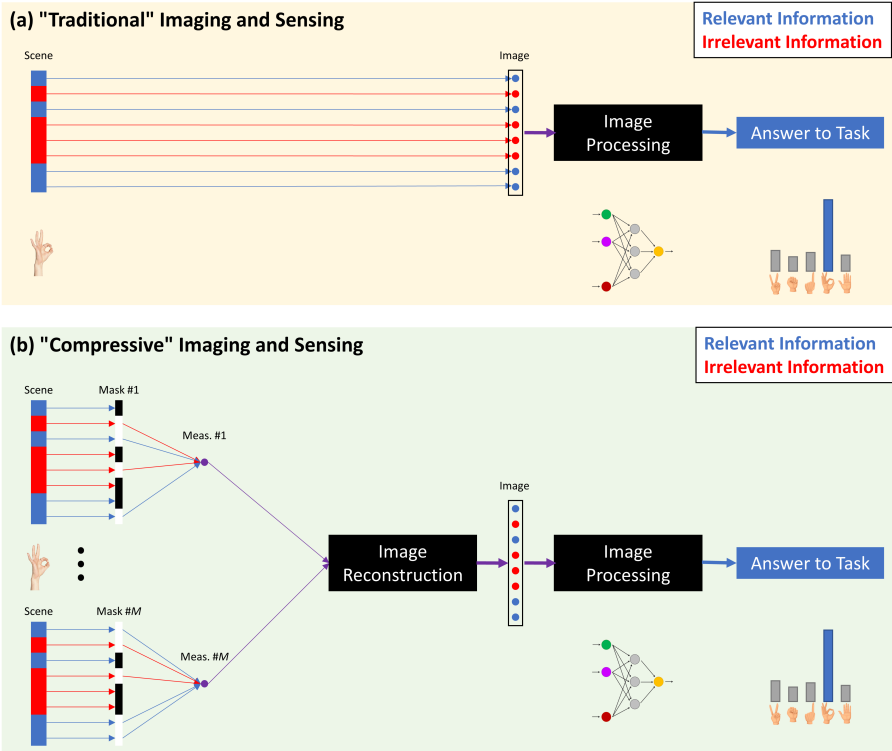


Figure 1. Comparison of the entire imaging pipeline for “traditional” and “compressive” imagers, from the measurement process all the way to answering a specific task. Both paradigms indiscriminately acquire both relevant (blue) and irrelevant (red) scene information and only separate the two in the very last step. The compressive imager has an additional image reconstruction step because it acquires multiplexed measurements.

A traditional imager is schematically depicted in Fig. 1(a). It is based on a one-to-one mapping between scene and image, and then the image is interpreted to identify the sought-after information. In other words, the separation of relevant and irrelevant information occurs only at the very last step. A “compressive” imager is illustrated in Fig. 1(b). Instead of the one-to-one mapping between scene and measurements, spatial scene information is multiplexed across a diverse set of measurement modes with a coded aperture. A series of measurements corresponding to different configurations of the coded aperture is taken, and an image of the scene is reconstructed and then interpreted. Assuming that the scene is sparse, this compressive scheme has the potential to reduce latency if the speed-up due to the lower number of measurements outweighs the additional

processing burden due to image reconstruction. However, again, relevant information is separated from irrelevant information only at the very last step.

Several ideas have been put forward to lower the latency of the compressive imager from Fig. 1(b). A first set of ideas revolves around reducing the processing burden. One proposal is to limit the image reconstruction to the most significant principal components of the measurement modes.² Another idea would be to skip the image reconstruction altogether and directly process the “raw” data. A second set of ideas revolves around the coded aperture configurations. Initially, a series of random configurations was used.³ Subsequent efforts aimed at achieving such randomness by leveraging complex scattering media, using their spatial,⁴ spectral⁵ or configurational⁶ diversity, or by fabricating man-made disordered meta-structures to the same effect.⁷⁻¹⁰ Random measurement modes are “quasi-orthogonal” but not truly orthogonal because they always have a finite overlap, resulting in an inevitable acquisition of redundant information in subsequent measurements. Hence, if truly orthogonal measurement modes could be used, no redundant information would be acquired such that fewer measurements would be needed. The implementation of truly orthogonal coded aperture configurations has been achieved by carefully configuring a reflect-array with patterns from the Hadamard basis¹¹ or by judiciously tuning the scattering properties of a complex medium.^{12,13}

All imagers discussed thus far are completely oblivious to scene, measurement hardware and task. In other words, they ignore available *a priori* knowledge such that these imagers are slowed down by the flow of irrelevant information through their measurement and processing layers. It is important to note that in a typical imaging scenario, vast amounts of *a priori* knowledge are available, including the expected scene (e.g. a human in security screening), the measurement hardware (e.g. a specific programmable-metasurface device) and the task (e.g. concealed weapon detection in security screening). The more use of this *a priori* knowledge is made, the more specific the imaging setup can be, and it is this tailoring that will enable the separation between relevant and irrelevant information earlier on in the pipeline.

A first notable effort in that spirit was the idea to make use of knowledge about the expected scene by using principle scene components as illumination patterns.^{14,15} These studies indeed achieved a reduction of latency by reducing the number of required measurements. However, they still ignored large portions of the available *a priori* knowledge, notably about the task. The measurement process did not discriminate between relevant and irrelevant information for the task at hand, implying the acquisition of useless (irrelevant) information.

Here, we describe a class of “intelligent” imagers that select task-relevant information already in the measurement process (see Fig. 2), thereby achieving significant latency improvements. The underlying principle (illustrated in Fig. 3) is to treat the entire pipeline from measurement to answering a specific task in an end-to-end fashion. Specifically, inspired by recent work on intelligent microscopes,¹⁶ we interpret reconfigurable measurement settings (in our case *in situ* programmable meta-atoms) as trainable physical weights that can be optimized in a unique pipeline together with the digital weights of an artificial neural network (ANN) on the processing side.¹⁷

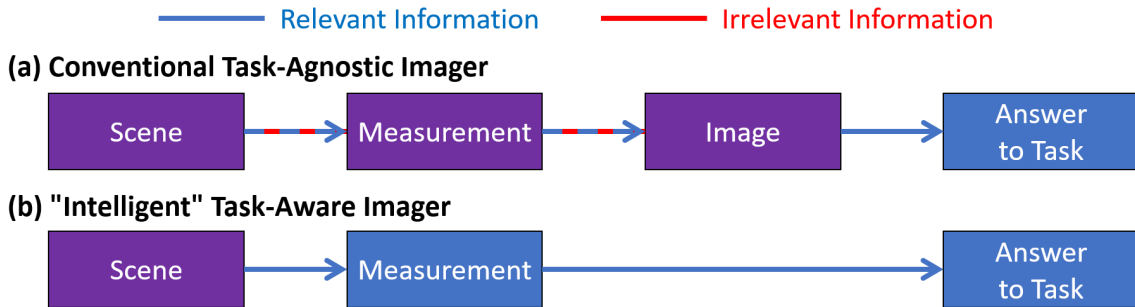


Figure 2. Conventional imagers (a) are *task-agnostic*, and drag irrelevant information through the entire pipeline until they discriminate between relevant and irrelevant information at the very last step. An “intelligent” imager (b) is *task-aware*, and makes this discrimination already during the measurement process. Moreover, the image reconstruction step is skipped and instead the raw data is directly interpreted.

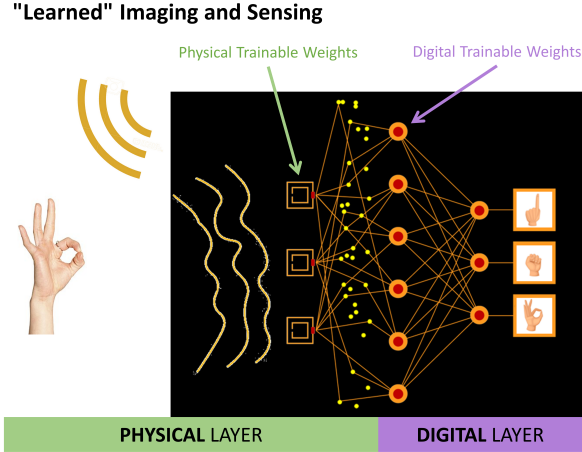


Figure 3. An artist’s impression of “Learned Sensing”. Information flows through a hybrid analog-digital ANN containing both physical and digital weights that are jointly trained. Thereby, the measurement process is aware of the ultimate task and discriminates between relevant and irrelevant information.

3. LEARNED SENSING

3.1 Setting up the hybrid analog-digital end-to-end pipeline

The first step toward implementing a “learned” sensing pipeline is to obtain a forward model of the physical measurement process in order to integrate it into the pipeline. In some scenarios, such a forward model can be of analytical nature. For instance, for the aforementioned planar dynamic metasurface transceivers, an efficient compact description of the radiating sub-wavelength meta-atoms as dipoles is feasible.¹⁷ In other scenarios, an accurate analytical description of the measurement process may be too cumbersome or impossible. In these cases, an ANN can be trained to approximate the forward model.¹⁸ A hybrid approach is also conceivable, in which due to fabrication inaccuracies a few parameters of the analytical model are fine-tuned based on experimental measurements. If an analytical forward model involves mathematical functions that are not easily differentiable, it may be preferable to resort to an ANN-based approximation of the forward model in light of the subsequent training phase.

The hybrid analog-digital end-to-end pipeline is then set-up in a suitable programming environment such as TensorFlow. The first layer is not a standard ANN-layer but a custom layer, namely the forward model of the measurement process.¹⁷ The programmable parameters therein, in our case the configuration of the programmable meta-atoms, are treated as the trainable variables of this first (physical) layer.¹⁷ From the first layer, the information seamlessly flows into subsequent layers that are conventional ANN-layers (e.g. fully connected layers), as seen in Fig. 3. Each of these digital layers, of course, also contains numerous trainable (digital) weights.

An interesting aside is the question what type of digital ANN-layers would be most suitable. Currently, convolutional architectures are very popular.¹⁶ Their popularity originates from their use in computer vision, where they are applied to “traditional” images as in Fig. 1(a) and excel at identifying relevant local correlations in the data. However, given that in our case the data is completely scrambled by the multiplexing (and we intend to skip the image reconstruction step), the sought information is very likely encoded in long-range correlations within the data such that fully connected layers appear more suitable.^{19–21}

3.2 Jointly training physical and digital weights

Once the hybrid pipeline is set up, the challenge lies in jointly training both physical and digital weights. Supervised learning requires access to a labelled training data set. Such data can be generated, for instance, based on binarized optical images of the scene.¹⁸ The optimization of the trainable weights is then usually accomplished via error backpropagation,²² hence the need for differentiable functions in the forward model. In many “learned

sensing” problems, the one-bit programmability of the meta-atoms implies binary (thus discontinuous) physical weight variables which are in principle incompatible with error backpropagation. However, various tricks such as the use of a “temperature parameter” can overcome this challenge.²³

It is important to note that the outcome of the training process depends to some extent on the random initialization of the variables. This is not surprising but in fact very common in most inverse problems. There is no guarantee of identifying the globally optimal set of configurations, but typically different optimization runs identify distinct sets of locally optimal configurations that perform roughly equally well. An interesting interpretation of this is to draw an analogy with statistical physics:²⁴ different configurations of the velocity of the air molecules correspond to the same room temperature, pressure, etc. In our problem, there are different realizations of the weights that correspond to roughly similar performance in terms of a specific task.

3.3 Performance benchmarking against the state-of-the-art

To evidence the substantial latency improvements with “learned sensing”, we benchmark the performance characteristics against the three previously discussed state-of-the-art scene illumination strategies (random patterns, orthogonal patterns, principal scene components) in the same imaging/sensing setup. Specifically, we consider in Fig. 4 the dependence of the achievable accuracy in a prototypical classification task as a function of the number of utilized scene illuminations. The fewer measurements are needed to achieve a targeted accuracy, the faster both the measurement and the processing are, resulting in a double advantage in terms of latency.

In Fig. 4(a) we consider a prototypical object recognition task in simulation. The scene contains a metallic digit that is to be recognized by illuminating the scene with a series of patterns formed by a planar dynamic metasurface transceiver.¹⁷ If very few measurements are allowed, our approach outperforms all three conventional techniques by up to 15 % in terms of achievable accuracy. Naturally, as more and more measurements are allowed, the gap gets narrower, and if many measurements are taken, no difference between the schemes is expected.

The results in Fig. 4(a) can also be discussed in terms of the amount of useful information obtained per measurement. A random set of illuminations yields the worst performance because it makes no attempt at discriminating between relevant and irrelevant information, and moreover acquires redundant information in subsequent measurements. Orthogonal patterns perform slightly better because they remove the redundancy issue but still fail to discriminate between relevant and irrelevant. A more notable improvement is obtained with principle scene components as illumination patterns because knowledge about the expected scene is used; yet knowledge about the task is still ignored. If knowledge about the task is deployed via the “learned sensing” paradigm, a remarkably better performance is achieved.

In Fig. 4(b) we consider a human gesture recognition task in experiments. This time, the scene is illuminate by waves reflected of a programmable metasurface reflect-array.¹⁸ In line with the previous case study, “learned sensing” enables higher classification accuracies with fewer measurements in comparison to the state-of-the art.

3.4 Can we understand why the learned patterns are optimal?

Having observed the superior performance of the learned patterns, one question remains: can we gain an intuitive understanding of them? To shed light on this question, we return to our case study from Fig. 4(a). Specifically, we compare the four illumination techniques in terms of the patterns’ overlap in Fig. 5(a) and the patterns’ resemblance to principle scene components in Fig. 5(b).¹⁷

First, we note in Fig. 5(a) that the overlap of the orthogonal patterns is close to zero, confirming that they were properly synthesized. The overlap of the learned patterns is similar to that of the random patterns, indicating that redundancy considerations do not play a major role. We further observe that the PCA patterns overlap only slightly less than the random patterns even though ideal PCA patterns are orthogonal by definition – hinting at the difficulty to properly synthesize PCA patterns under the given hardware constraints. This is also evidenced in Fig. 5(b) which reveals that the overlap of the PCA-based patterns with the principal scene components is high but far from unity. The learned patterns have slightly more overlap with the principle scene components than random or orthogonal patterns which are both scene-agnostic. This explains why solely making use of knowledge about the expected scene as in Refs.^{14,15} only offers limited latency gains well below those achievable with “learned sensing”: the optimized learned patterns only mildly take into account the scene

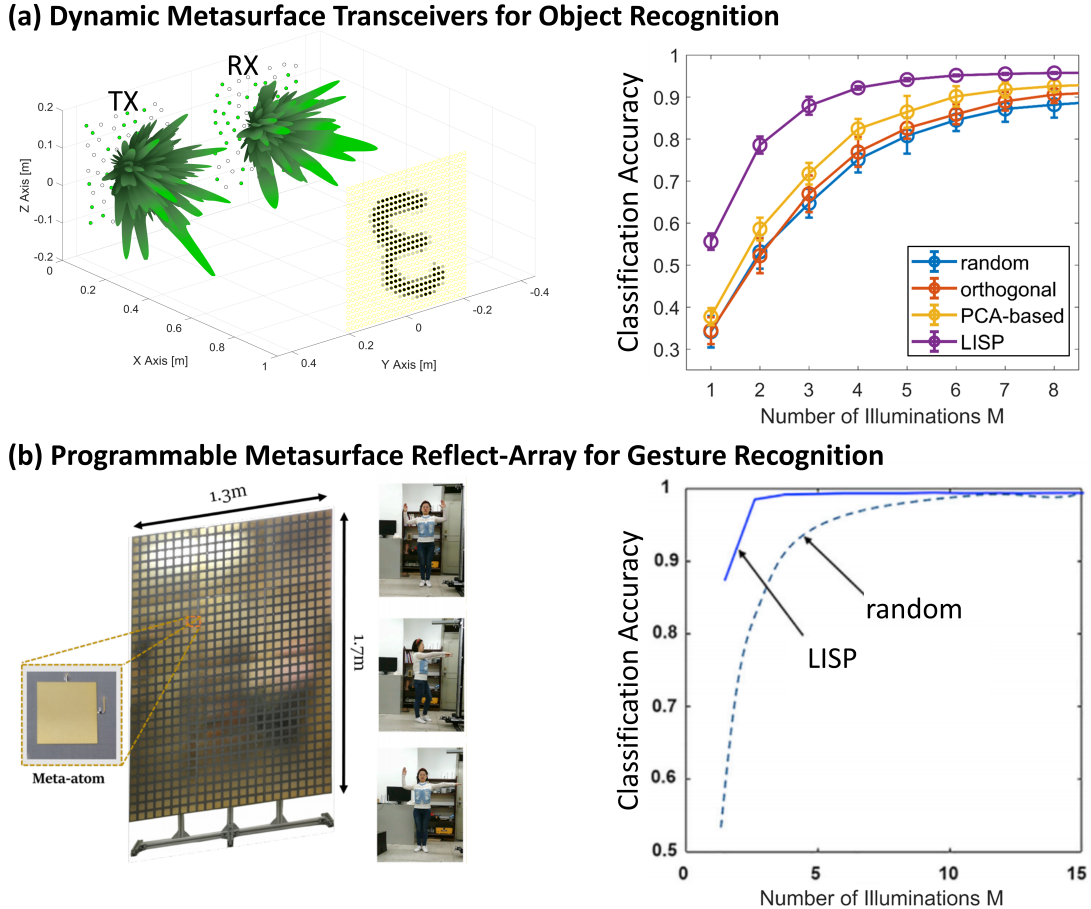


Figure 4. Two implementations of learned sensing with programmable metasurface hardware. (a) Numerical study considering planar dynamic metasurface transceivers as hardware for object (metallic digit) recognition.¹⁷ The dependence of the object classification accuracy on the number of utilized scene illuminations is plotted (“LISP”) and compared to three state-of-the-art approaches (random, orthogonal, principal scene components). (b) Experiment based on a programmable metasurface reflect-array as hardware for human gesture recognition.¹⁸ The dependence of the gesture classification accuracy on the number of utilized scene illuminations is plotted and compared against the use of random scene illuminations.

knowledge. Overall, Fig. 5 demonstrates that the learned patterns are not easily explained in terms of overlap or principal scene components because they take into account the specific task; they are hence not intuitive and can only be obtained through the end-to-end hybrid analog-digital pipeline discussed above.

4. CONCLUSION AND OUTLOOK

In summary, we have discussed the evolution from “compressed” to “learned” sensing from an information flow perspective. Given the quest to reduce the latency of electromagnetic imagers as much as possible, we showed that discriminating between relevant and irrelevant information already during the measurement process rather than at the very last processing step is crucial. We outlined a general recipe to that effect which consists of a hybrid analog-digital pipeline containing trainable physical and digital weights, encompassing both the programmable measurement process as well as the data processing. Jointly optimizing physical and digital weights allows the measurement process to be task-aware by selecting relevant information. Thereby, more useful information is obtained per measurement such that overall latency is drastically improved. We illustrated this for two prototypical recognition tasks with two distinct examples of single-feed programmable-metasurface hardware.

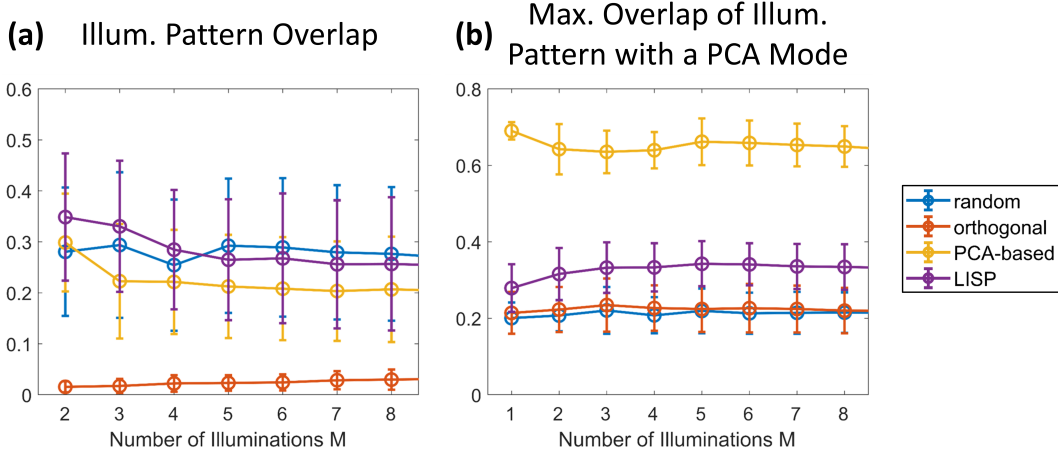


Figure 5. Analysis of the learned illumination patterns from Fig. 3(a) in terms of their overlap (a) and their resemblance to principle scene components (b).

While this paper has focused on latency, it is worthwhile noting that limiting the number of necessary measurements is also favorable with respect to numerous further metrics, including power consumption, radiation exposure, required computational resources and cost. Moreover, although the considered programmable metasurface devices appear advantageous in comparison to traditional antenna arrays in terms of the hardware complexity, the “learned sensing” paradigm is equally applicable to antenna array architectures such as phased arrays. The trainable physical parameters in that case could be the phase and/or amplitude settings of each antenna.

Looking forward, the information flow perspective may offer further opportunities for reducing latency in electromagnetic imaging. Specifically, a dynamic (rather than static) sequence of coded aperture configurations that in addition takes into account information from preceding measurements is an enticing avenue for future investigations,²⁵ akin to the recurrent visual attention of humans.²⁶ Within the context of metasurface-based hardware, such a step may provide the programmable meta-atoms with a “cognitive” feature.

A second topic for future research is the application of the “learned sensing” paradigm to a special type of coded aperture: the “reverberation coded aperture” (RCA).²⁷ RCAs are coded apertures based on complex media for which the scene is *within* rather than *outside* the complex medium. This difference means that waves interact with the scene not once but countless times, hence developing a much stronger sensitivity to sub-wavelength scene details. There is a direct link between the waves’ dwell time in the complex medium, their sensitivity to sub-wavelength details and the achievable resolution.²⁷ Without any manipulation of the scene’s near field, RCAs enable the recovery of deeply sub-wavelength details – a fact that does not violate the diffraction limit because this limit ignores available *a priori* knowledge, the signal-to-noise ratio and the fact that coded aperture imagers are not even based on focusing. By optimizing an RCA’s scattering properties in a “learned sensing” framework, further resolution improvements appear to be achievable. The practical relevance of RCAs lies in the fact that many electromagnetic imaging scenarios naturally involve a scene enclosed by a scattering enclosure, for instance, imaging and sensing in an indoor environment, a metro station or an airplane.

REFERENCES

- [1] Donoho, D. L., “Compressed sensing,” *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006).
- [2] Fromentèze, T., Yurduseven, O., del Hougne, P., and Smith, D. R., “Lowering latency and processing burden in computational imaging through dimensionality reduction of the sensing matrix,” *Sci. Rep.* **11**(1), 1–14 (2021).
- [3] Chan, W. L., Charan, K., Takhar, D., Kelly, K. F., Baraniuk, R. G., and Mittleman, D. M., “A single-pixel terahertz imaging system based on compressed sensing,” *Appl. Phys. Lett.* **93**(12), 121105 (2008).

- [4] Liutkus, A., Martina, D., Popoff, S., Chardon, G., Katz, O., Lerosey, G., Gigan, S., Daudet, L., and Carron, I., “Imaging with nature: Compressive imaging using a multiply scattering medium,” *Sci. Rep.* **4**, 5552 (2014).
- [5] Fromenteze, T., Yurduseven, O., Imani, M. F., Gollub, J., Decroze, C., Carsenat, D., and Smith, D. R., “Computational imaging using a mode-mixing cavity at microwave frequencies,” *Appl. Phys. Lett.* **106**(19), 194104 (2015).
- [6] Sleasman, T., Imani, M. F., Gollub, J. N., and Smith, D. R., “Microwave imaging using a disordered cavity with a dynamically tunable impedance surface,” *Phys. Rev. Applied* **6**(5), 054019 (2016).
- [7] Hunt, J., Driscoll, T., Mrozack, A., Lipworth, G., Reynolds, M., Brady, D., and Smith, D. R., “Metamaterial apertures for computational imaging,” *Science* **339**(6117), 310–313 (2013).
- [8] Xie, Y., Tsai, T.-H., Konneker, A., Popa, B.-I., Brady, D. J., and Cummer, S. A., “Single-sensor multispeaker listening with acoustic metamaterials,” *Proc. Natl. Acad. Sci. U.S.A.* **112**(34), 10595–10598 (2015).
- [9] Sleasman, T., Imani, M., Gollub, J. N., and Smith, D. R., “Dynamic metamaterial aperture for microwave imaging,” *Appl. Phys. Lett.* **107**(20), 204104 (2015).
- [10] Sleasman, T., Imani, M. F., Diebold, A. V., Boyarsky, M., Trofatter, K. P., and Smith, D. R., “Implementation and characterization of a two-dimensional printed circuit dynamic metasurface aperture for computational microwave imaging,” *IEEE Trans. Antennas Propag.* (2020).
- [11] Watts, C. M., Shrekenhamer, D., Montoya, J., Lipworth, G., Hunt, J., Sleasman, T., Krishna, S., Smith, D. R., and Padilla, W. J., “Terahertz compressive imaging with metamaterial spatial light modulators,” *Nat. Photonics* **8**(8), 605–609 (2014).
- [12] del Hougne, P., Fink, M., and Lerosey, G., “Optimally diverse communication channels in disordered environments with tuned randomness,” *Nat. Electron.* **2**(1), 36–41 (2019).
- [13] del Hougne, P., Davy, M., and Kuhl, U., “Optimal multiplexing of spatially encoded information across custom-tailored configurations of a metasurface-tunable chaotic cavity,” *Phys. Rev. Applied* **13**(4), 041004 (2020).
- [14] Liang, M., Li, Y., Meng, H., Neifeld, M. A., and Xin, H., “Reconfigurable array design to realize principal component analysis (pca)-based microwave compressive sensing imaging system,” *IEEE Antennas Wirel. Propag. Lett.* **14**, 1039–1042 (2015).
- [15] Li, L., Ruan, H., Liu, C., Li, Y., Shuang, Y., Alù, A., Qiu, C.-W., and Cui, T. J., “Machine-learning reprogrammable metasurface imager,” *Nat. Commun.* **10**(1), 1–8 (2019).
- [16] Horstmeyer, R., Chen, R. Y., Kappes, B., and Judkewitz, B., “Convolutional neural networks that teach microscopes how to image,” *arXiv preprint arXiv:1709.07223* (2017).
- [17] del Hougne, P., Imani, M. F., Diebold, A. V., Horstmeyer, R., and Smith, D. R., “Learned integrated sensing pipeline: Reconfigurable metasurface transceivers as trainable physical layer in an artificial neural network,” *Adv. Sci.* **7**(3), 1901913 (2020).
- [18] Li, H.-Y., Zhao, H.-T., Wei, M.-L., Ruan, H.-X., Shuang, Y., Cui, T. J., del Hougne, P., and Li, L., “Intelligent electromagnetic sensing with learnable data acquisition and processing,” *Patterns* **1**(1), 100006 (2020).
- [19] Zheng, Y., Huang, J., Chen, T., Ou, Y., and Zhou, W., “Processing global and local features in convolutional neural network (cnn) and primate visual systems,” in [*Mobile Multimedia/Image Processing, Security, and Applications 2018*], **10668**, 1066809, International Society for Optics and Photonics (2018).
- [20] del Hougne, P., “Robust position sensing with wave fingerprints in dynamic complex propagation environments,” *Phys. Rev. Research* **2**, 043224 (2020).
- [21] Zhu, C., Chan, E. A., Wang, Y., Peng, W., Guo, R., Zhang, B., Soci, C., and Chong, Y., “Image reconstruction through a multimode fiber with a simple neural network architecture,” *Sci. Rep.* **11**, 896 (2021).
- [22] LeCun, Y., Bengio, Y., and Hinton, G., “Deep learning,” *Nature* **521**(7553), 436–444 (2015).
- [23] Chakrabarti, A., “Learning sensor multiplexing design through back-propagation,” in [*Proceedings of the 30th International Conference on Neural Information Processing Systems*], *NIPS’16*, 3089–3097, Curran Associates Inc., Red Hook, NY, USA (2016).
- [24] Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., and Zdeborová, L., “Machine learning and the physical sciences,” *Rev. Mod. Phys.* **91**, 045002 (2019).

- [25] Chaware, A., Cooke, C. L., Kim, K., and Horstmeyer, R., “Towards an intelligent microscope: adaptively learned illumination for optimal sample classification,” in [*ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*], 9284–9288, IEEE (2020).
- [26] Mnih, V., Heess, N., Graves, A., and Kavukcuoglu, K., “Recurrent models of visual attention,” in [*Advances in Neural Information Processing Systems*], **27**, Curran Associates, Inc. (2014).
- [27] del Hougne, M., Gigan, S., and del Hougne, P., “Deeply sub-wavelength localization with reverberation-coded-aperture,” *arXiv preprint arXiv:2102.05642* (2021).