

Reconstruction-free Inference on Compressive Measurements

Suhas Lohit¹, Kuldeep Kulkarni^{1,2}, Pavan Turaga^{1,2}, Jian Wang³, and Aswin C. Sankaranarayanan³

¹School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ

²School of Arts, Media and Engineering, Arizona State University, Tempe, AZ

³Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA

Abstract

Spatial-multiplexing cameras have emerged as a promising alternative to classical imaging devices, often enabling acquisition of ‘more for less’. One popular architecture for spatial multiplexing is the single-pixel camera (SPC), which acquires coded measurements of the scene with pseudo-random spatial masks. Significant theoretical developments over the past few years provide a means for reconstruction of the original imagery from coded measurements at sub-Nyquist sampling rates. Yet, accurate reconstruction generally requires high measurement rates and high signal-to-noise ratios. In this paper, we enquire if one can perform high-level visual inference problems (e.g. face recognition or action recognition) from compressive cameras without the need for image reconstruction. This is an interesting question since in many practical scenarios, our goals extend beyond image reconstruction. However, most inference tasks often require non-linear features and it is not clear how to extract such features directly from compressed measurements. In this paper, we show that one can extract non-trivial correlational features directly without reconstruction of the imagery. As a specific example, we consider the problem of face recognition beyond the visible spectrum e.g. in the short-wave infra-red region (SWIR) – where pixels are expensive. We base our framework on smashed filters which suggests that inner-products between high-dimensional signals can be computed in the compressive domain to a high degree of accuracy. We collect a new face image dataset of 30 subjects, obtained using an SPC. Using face recognition as an example, we show that one can indeed perform reconstruction-free inference with a very small loss of accuracy at very high compression ratios of 100 and more.

1. Introduction

Compressed sensing [5] provides a way of sampling signals at sub-Nyquist rates. Assuming that the original signal is sparse, it can be recovered almost perfectly by using reconstruction algorithms. The single-pixel architecture [29] for compressive sensing presents a compelling use-case in spectral regions where the cost of a pixel may be prohibitively expensive – such as in the short-wave infrared (SWIR) region. While much work has focused on theory and algorithms for signal recovery, less attention has been devoted to whether one can perform high-level inference directly on compressive measurements without reconstruction. This question is motivated due to the following reasons: a) in many practical scenarios one is often interested in inferring some property of the scene rather than knowing the entire scene itself, b) high fidelity reconstruction is difficult at high compression ratios, c) reconstruction quality is sensitive to several parameters such as signal sparsity, sparsifying basis, signal-to-noise ratio etc, many of which are chosen rather arbitrarily in a real application.

In this paper, we hypothesize that the reconstruction problem can be entirely bypassed when a specific inference task is the eventual goal of sensing. Instead of reconstruction, we ask what class of features could one reliably extract from compressive cameras which provides robust high-level inference capabilities. Specifically, we show that one can extract *correlational features* directly from compressive measurements. Correlational features provide a solid foundation to devise a number of high-level inference algorithms, and have been used widely in computer vision literature [14].

As a specific example, we focus on the problem of face recognition in the Near-infrared (NIR) spectrum from compressively sensed measurements of the face. Infrared imaging has become an attractive sensing modality for face recognition. The lack of infrared energy in indoor illumination allows for precise control of lighting in infrared

The work of SL, KK and PT was supported by ONR Grant N00014-12-1-0124 subaward Z868302. ACS and JW were supported in part by the NSF Grant CCF 1117939.

and hence, provides the ability to isolate illumination variations [15]². However, infrared cameras are very expensive, and this has prevented them from being employed for tasks like face recognition. The single-pixel camera (SPC) architecture [29] provides a cost-effective solution for the acquisition problem. The SPC employs a single photodiode and a micro-mirror array to acquire images. This greatly reduces the cost of the camera as a single photodetector, sensitive to wavelengths of interest, is used for data acquisition.

1.1. Related work

Compressive Sensing: It was proven by Candes *et al.* [4] and Donoho [9], that for a signal $\mathbf{x}$ of dimension n which is s -sparse in some basis Ψ , it is possible to sense it at a rate much lower than the Nyquist rate and still be able to recover it almost perfectly. As opposed to samples of a signal, the workhorse underlying the celebrated Nyquist theorem, CS suggests obtaining non-adaptive linear measurements of the signal. The number of such measurements required for near-perfect reconstruction is shown to be $O(s \log(\frac{n}{s}))$. The SPC, developed for obtaining arbitrary linear transformations of a scene, uses a micro-mirror array to encode the measurement matrix Φ of dimension $m \times n$, $m < n$ to obtain measurements of the form $\mathbf{y} = \Phi\mathbf{x}$. It is required that the measurement basis Φ be incoherent with the sparsifying basis Ψ [5]. Recovering $\mathbf{x}$ from $\mathbf{y}$ forms an underdetermined linear system but, when the sparsity condition is satisfied, it becomes possible to recover $\mathbf{x}$. Many algorithms such as Orthogonal Matching Pursuit [28] and Basis Pursuit [6] have been proposed for this task but they suffer from high computational complexity, which makes them impractical to use in scenarios where inference is the eventual goal. A great deal of research has been devoted to the development of faster reconstruction algorithms. In this paper, we answer two related questions, 1) If the final goal is only inference, e.g. the identity of a face, and not reconstruction of the image itself, is it possible to achieve it directly from compressed measurements thus avoiding reconstruction? and 2) Assuming this is indeed possible, what methods employed for traditional face recognition with conventional cameras can be adapted to perform face recognition in the set up mentioned in 1) ?

Correlational features for inference: The concept of a correlation filter (CF) is a well researched idea and has been applied in target recognition, face verification etc. for over three decades [12]. The basic concept is to design a filter such that its correlation with a test image produces a large peak in the correlation output if it belongs to the true class and no significant peak otherwise. The main advantage of

using correlation filters is the shift invariance that it provides implicitly.

There is a large number of correlation filters available in the literature such as the Minimum Average Correlation Energy (MACE) [17] filter based on the idea of designing a filter that reduces the energy of the correlation output. It is shown that this leads to a large peak in the correlation plane at the target location. A more computationally efficient version of the MACE filter, called the unconstrained MACE (UMACE) was proposed by Mahalanobis *et al.* [19] which requires the inversion of a diagonal matrix which can be done easily. Although CFs provide very good localization, they do not perform as well as Support Vector Machines (SVMs) in terms of discrimination. To address this, more sophisticated correlation filters have been developed such as the Maximum Margin Correlation Filter (MMCF) [23]. MMCFs combine the strengths of CFs and SVMs and thus improve the generalization capability of the filter while maintaining localization.

Later in this paper, we conduct experiments on face recognition; so we review some relevant work in this area here. An established method of performing face recognition is by first extracting features from face images and then using pattern recognition techniques for recognition. These features include linear features such as Gabor features [16], Principal Component Analysis (PCA) [26], Linear Discriminant Analysis (LDA) [10] etc. and non-linear features such as Histogram of Oriented Gradients (HOG) [8] and Local Binary Patterns (LBP) [1] and combinations of these features. Recently, deep learning and convolutional neural networks have been employed [27] on very large datasets to achieve very good recognition rates. *At present, it is unclear how to derive non-linear features from compressed measurements.*

Inference from compressive measurements: Calderbank *et al.* [3] show that classifiers can be designed directly in the compressed domain. However, they consider classifiers to be learnt from the raw data, and do not consider the role of feature-extraction, which is the core of this paper. Davenport *et al.* [7] propose the idea of the smashed filter to perform classification directly on the compressed measurements. Specifically, by invoking the Johnson-Lindenstrauss lemma [11], they show that likelihoods can be meaningfully computed from compressive measurements. In [24], Sankaranarayanan *et al.* develop a framework for acquiring CS videos and their reconstruction by modeling videos as linear dynamical systems (LDS). They show preliminary results on classification performed on LDS parameters obtained directly from compressive measurements. Neifeld and Shankar [22] propose ‘feature-specific imaging’ where images are directly measured in the required task-specific basis such as Karhunen-Loeve or wavelet basis. Baheti and

²The same principle is used in Kinect to obtain depth images via structured illumination in NIR

Neifeld [2] develop an adaptive version of feature-specific imaging where knowledge about past measurements are used to design an optimal projection basis for future measurements. They apply this framework to face recognition and report superior results. Their work is different from this paper since we obtain data-independent measurements using a random measurement matrix and extract discriminative features from these measurements directly. In [18], a compressed sensing architecture is developed where, instead of perfect reconstruction of the CS images, only relevant parts of the scene i.e., the objects are reconstructed. In [30] and [20], ideas from compressive sensing are used in face recognition with very good results. They rely on finding a sparse code for the test image in terms of the training set vectors which is analogous to reconstruction in compressed sensing. A closely related recent work by Kulkarni *et al.* [13] tackles the problem of reconstruction-free action recognition from compressive cameras. Face recognition in the framework of reconstruction free inference from compressively sensed images has not been attempted to the best of our knowledge.

1.2. Contributions

1. In this paper, we propose a framework for extraction of correlational features on compressive measurements sensing without having to perform reconstruction.
2. We show that this framework performs well at very high compression ratios where reconstruction would otherwise fail.
3. We show that one can perform inference of faces quite robustly with only a small loss in accuracy compared to oracle sensing.

The paper is organized as follows. In section 2, the algorithm for reconstruction-free face recognition is described in detail. In section 3, experiments are performed on visible and near infrared face datasets as well as actual compressed measurements. We present conclusions and directions for future work in section 4.

2. Correlations as a basis for inference

Many high-level problems such as scene, action, and face recognition are founded on the basis of computing certain correlational features. Examples include generic filters such as Gabor or wavelet banks as well as purposively trained filters such as MACH or MACE filters [17, 19]. We ask the following question: what kind of filters are amenable to robust inference, in our case for face recognition. We employ MMCF which provides SVM-like max-margin properties while retaining the correlation structure that, as we will see soon, is invaluable in the context of CS. If one has access to a set of filters which have been deemed

useful for a given inference problem, can we then compute the filter-responses directly, without reconstruction, on the compressive measurements? This question leads us to employ the technique of smashed filters [7]. We review this technique in section 2.2. The block diagram of the proposed method is shown in Figure 1.

2.1. Training MMCFs

Consider a training set consisting of face images represented as column vectors $\mathbf{x}_i$, $i = 1, 2, \dots, N$ belonging to two classes (when we have more than two classes, we use a one-versus-all training strategy). The MMCF [23] attempts to find a filter $\mathbf{w}$ with two key properties:

- (i) Low side-lobes: The inner-product of the filter to shifted versions of a signal needs to have low-energy. This criterion comes from earlier work in MACE filters.
- (ii) Max-margin: we seek to establish a large margin between the positive and negative training examples. This, the objective of SVM, reduces to correlations at the center of the image being well-separated.

Both of these properties can be achieved by solving an optimization of the form:

$$\begin{aligned} \min_{\mathbf{w}, b} \quad & (\mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i, \sum_{i=1}^N \|\mathbf{w} \otimes \mathbf{x}_i - \mathbf{g}_i\|_2^2) \quad (1) \\ \text{s.t.} \quad & t_i(\mathbf{w}^T \mathbf{x}_i + b) \geq c_i - \xi_i, \quad i = 1, 2, \dots, N, \end{aligned}$$

where the minimizer $\mathbf{w}^*$ is the required MMCF as well as the separating hyperplane, $\mathbf{g}_i$ is the desired value of the correlation output, $c_i = 1$ for images in the true class and $c_i = 0$ for those in the false class. ξ_i are the positive slack variables that take care of outliers and $t_i \in \{-1, 1\}$ are the labels. As shown in [23], with appropriate transformations, (1) can be reduced to a single optimization problem that can be solved on any standard SVM solver.

2.2. Compressive smashed filtering

Once the correlation filters are obtained for each of the classes, they can now be used on the testing set as follows. Let X be a non-compressed input image of size $N_1 \times N_2$. Let H_m be the correlation filter for the m^{th} class, $m = 1, 2, \dots, M$. H is also of size $N_1 \times N_2$. In the traditional sensing framework, the input image is to be correlated with each of these filters to obtain the correlation plane c of size $(2N_1 - 1) \times (2N_2 - 1)$ given by

$$c_m(i, j) = \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} X(k, l) H_m(k - i, l - j). \quad (2)$$

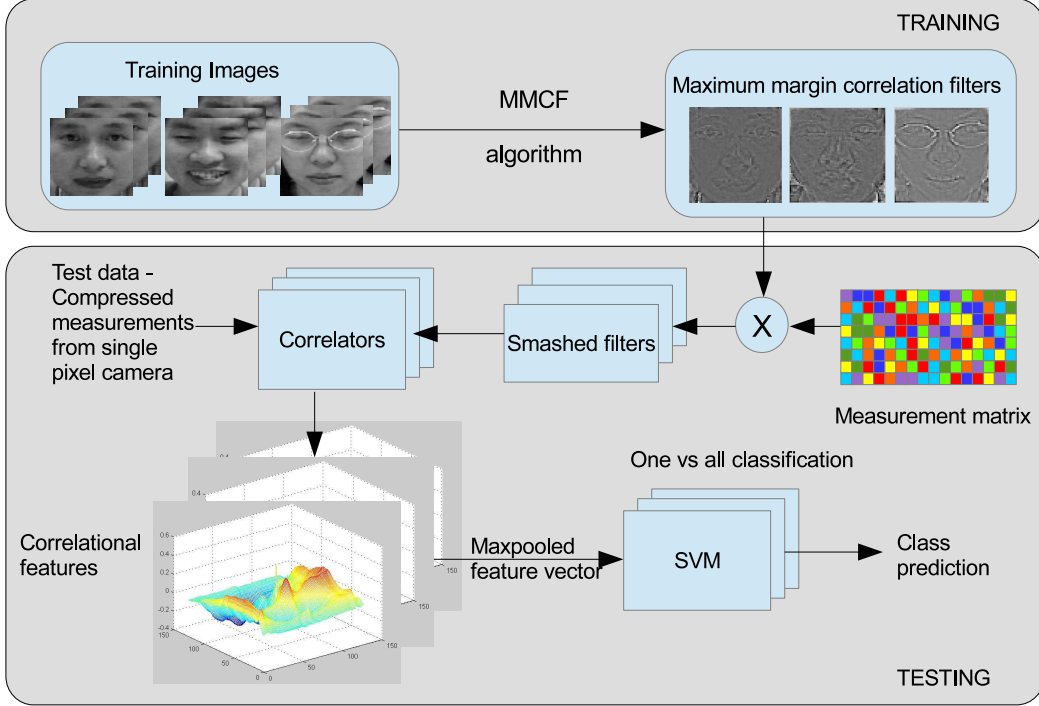


Figure 1. Illustration of reconstruction-free correlational feature estimation for high-level visual inference problems, exemplified by face recognition in this paper.

This can be written in the form of an inner product as $c_m(i, j) = \langle X, H_m^{i,j} \rangle$, where $H_m^{i,j}$ is the shifted version of H_m by i and j units in the x and y directions respectively. The measurement matrix, Φ , is chosen such that it preserves distances between points to a certain degree of accuracy when projecting from a high dimensional to a low dimensional space. More formally, using the Johnson-Lindenstrauss lemma [13], it can be shown:

$$c_m(i, j) - \epsilon \leq \langle \Phi X, \Phi H_m^{i,j} \rangle \leq c_m(i, j) + \epsilon, \quad (3)$$

where ΦX is the output from the single pixel camera, $\Phi H_m^{i,j}$ is the smashed filter and $\epsilon > 0$ [7]. Thus, $c_m(i, j)$ can be approximated by $\hat{c}_m(i, j)$ as

$$\hat{c}_m(i, j) = \langle \Phi X, \Phi H_m^{i,j} \rangle \quad (4)$$

This means that the correlation outputs of the compressed measurements can be obtained to a certain degree of accuracy (determined by the number of measurements) without reconstruction. Clearly, if there are M subjects, M correlation filters are trained and M correlation planes are obtained for each test image.

Extracting the feature vector from correlation planes: Each correlation plane is divided into non-overlapping

blocks and for each block, the peak and peak to side-lobe ratio (PSR) are determined. PSR is calculated using the formula $PSR = \frac{peak - \mu}{\sigma}$, where μ is the mean and σ is the standard deviation of the correlation values in a bigger region around a mask centered at the peak as explained in [25]. The peaks and PSRs of the different blocks are concatenated. Similar vectors are obtained for the each of the M correlation planes. All these vectors are concatenated to form a single feature vector for the particular test image. This feature vector is input into M linear SVMs for a one vs all classification. It is to be noted that the SVMs are trained on feature vectors obtained in the same fashion from the training set.

3. Experimental Results

The framework described in Section 2 is applied on different datasets to show that face recognition can be performed directly on the compressed measurements obtained from the SPC without reconstruction. We conduct two sets of experiments – the first is a controlled set of experiments on publicly available databases, where we simulate compressive acquisition in software. In the second set, we use real data from a compressive sensing architecture. In the controlled experiments, we experiment with Gaussian and Hadamard matrices for sensing, and show that inference is possible with only a small fraction of the total number of measurements in both cases, but Hadamard matrices

are more robust in the presence of measurement noise. In the experiments on real data, we only consider a permuted Hadamard measurement matrix.

3.1. Controlled Experiments

In this section, we conduct face recognition experiments on publicly available databases and simulate compressed sensing in software. We also consider 3 kinds of measurement matrices – 1) random Gaussian Φ^G , 2) low-rank permuted Hadamard Φ^H , 3) a simple downsampling operator Φ^D .

NIR database: The NIR database [15] consists of near infrared images of 197 subjects with 20 images per subject. Each image was resized to 256×256 from the original size of 640×480 . For each subject, 10 images are used for training and the remaining 10 images are used for testing. Sample images from the dataset are shown in Figure 2.

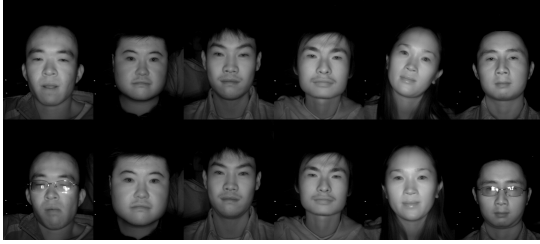


Figure 2. Sample images from the NIR database.

AMP database: The AMP database³ is a facial expression database compiled by the Advanced Multimedia Processing lab at CMU. The dataset consists of 975 grayscale facial images belonging to 13 people, each of size 64×64 pixels. Sample images from the database are shown in Figure 3. For each subject, 25 images are used for training and the remaining 50 images are used for testing.



Figure 3. Sample images from the AMP database.

Training and testing protocol: We here describe the protocol used for both datasets. For brevity, we have provided detailed parameters for the NIR database only where images are of resolution 256×256 . The AMP database uses corresponding scaled parameters adapted to the image-size 64×64 . For the NIR database, the MMCFs, one for each of the 197 classes, are trained using the 256×256 images from the test set as described in section 2.1. Then, each image in the dataset is vectorized to get a vector $\mathbf{x}_i$. The process of getting the measurements $\mathbf{y}_i$ from a single pixel camera is simulated using the equation $\mathbf{y}_i = \Phi \mathbf{x}_i$, where Φ is the measurement matrix.

First, the measurement matrix is chosen to be a Gaussian matrix (Φ^G) such that the entries of the matrix were i.i.d. standard Gaussian. The number of rows, r of the matrix corresponding to the number of compressed measurements is varied. Three values of r are chosen – 65536, 625 and 121 corresponding to compression ratios $CR = 1, 105, 542$ respectively.

Then, the trained correlation filters $\{H_i\}$ are also compressed to obtain the smashed filters $\tilde{H}_i = \Phi H_i$. As explained in section 2.2, each compressed image in the training set is correlated with all the smashed filters to obtain 197 correlation planes. Instead of using Equation 4, each correlation plane can be computed by first projecting the measured vector y_i back into the pixel space by pre-multiplying with Φ^T and then, correlating this result with the original filter H_m , since $\langle \Phi X, \Phi H_m^{i,j} \rangle = \langle \Phi^T \Phi X, H_m^{i,j} \rangle$. This can be efficiently computed using 2D-FFT by considering $\Phi^T \Phi X$ and H_m as the two signals.

Each correlation plane is divided into $B = 16$ square non-overlapping blocks (of size 128×128) and the PSR and peak values of each block is extracted. These values, in addition to the PSR and peak value for the entire correlation plane, are concatenated to form a feature vector of size 1×6698 . These features are used to train 197 linear SVMs, one for each class. In the testing phase, feature vectors of the compressed images are obtained in the same fashion and input to the trained SVMs for a one vs all classification. The accuracy of the face recognition system is determined as the ratio of number of correctly recognized faces to the total number of faces.

The above experiment is then repeated with Φ^H , the matrix containing a subset of rows of a permuted Hadamard matrix. The accuracy is determined for different numbers of measurements, r (the number of randomly chosen rows of Φ^H , as before. Finally, the images in the dataset are downsampled by the same factors.

Next, the effect of adding noise on face recognition accuracy is considered. Each of the above experiments is repeated after adding measurement noise – Gaussian noise – of standard deviation σ calculated using $\sigma = \eta \frac{\|\Phi \mathbf{x}\|}{\sqrt{M}}$, where $\eta = 0, 0.1, 0.2, 0.3$ is the noise level. The recognition ac-

³<http://chenlab.ece.cornell.edu/projects/FaceAuthentication/>

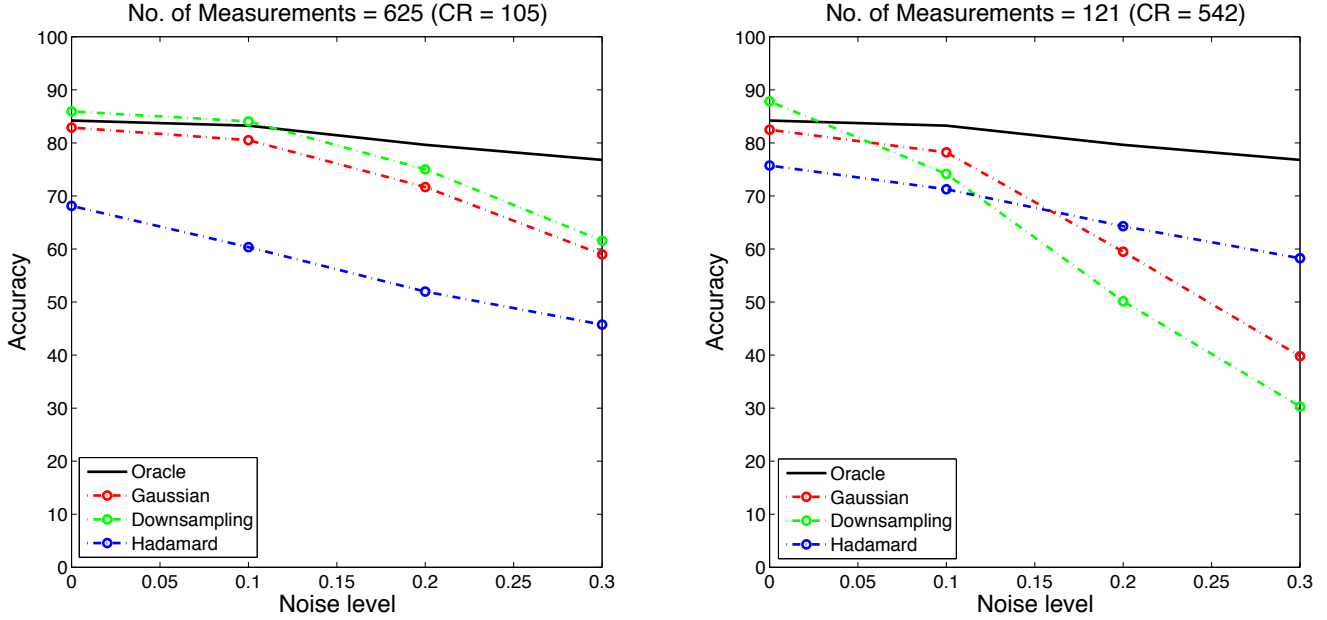


Figure 4. The figures show the variation of recognition accuracy for the NIR database for Oracle (no compression), Gaussian measurements, low-rank permuted Hadamard measurements, downsampling, for varying amounts of measurement noise. Note that results indicate that performance is close to Oracle for low-noise levels, and Hadamard is more stable in performance than Gaussian and downsampling operators.

curacies are determined at each noise level at each compression factor for each of the measurement matrices. The results are displayed in Figure 4.

Similar experiments are conducted on the AMP database at two compression ratios of 28 and 114. Since there are 13 subjects, 13 correlation filters are trained, each filter corresponding to one of the subjects. Each correlation plane is divided into $B = 4$ blocks and features are computed similar to above. The results are shown in Figure 5.

Salient observations: From the results shown in figures 4 and 5, we make the following observations:

1. *Reconstruction-free inference results in performance that closely matches Oracle (no compression) performance at low noise-levels.*
2. *Hadamard measurement matrices result in much more stable performance across noise compared to Gaussian or downsampling operators.*

3.2. Experiments on a single pixel camera

We built a single-pixel camera (SPC) using a DLP V7000 digital micromirror device (DMD), a Thorlabs PDA100A photodetector and a Digilent Analog Discovery

analog-to-digital converter. This is shown in Figure 6 The DMD has a resolution of 1024×768 and changes the micromirror configurations at a frame-rate of 22.7 kHz. The measurement rate of the SPC is determined primarily by the operating speed of the DMD; hence, we obtain 22.7k measurements per second.

Given a desired image resolution of $N \times N$ pixels, a Nyquist-camera would require N^2 measurements. This is reduced significantly when we invoke image priors and recovery techniques underlying CS. For example, to capture an image of resolution 128×128 without CS recovery, we would need 0.72 seconds at the operating rate of 22.7kHz. With CS, this can be reduced to as little as 0.1 seconds without significant loss in quality.

Based on the observations from the controlled experiments reported in section 3.1, we only use a permuted Hadamard as the measurement matrix. More specifically, for an $N \times N$ image, we first generate a $N^2 \times N^2$ column-permuted Hadamard matrix. Each row of this matrix is shaped into an $N \times N$ image that is upsampled and mapped to the 1024×768 micromirror array. Given that the DMD can only direct light towards or away from the photodetector, this implements a 0/1 measurement matrix. To obtain measurements corresponding to the ± 1 Hadamard matrix, we subtract half the average light level from the observed

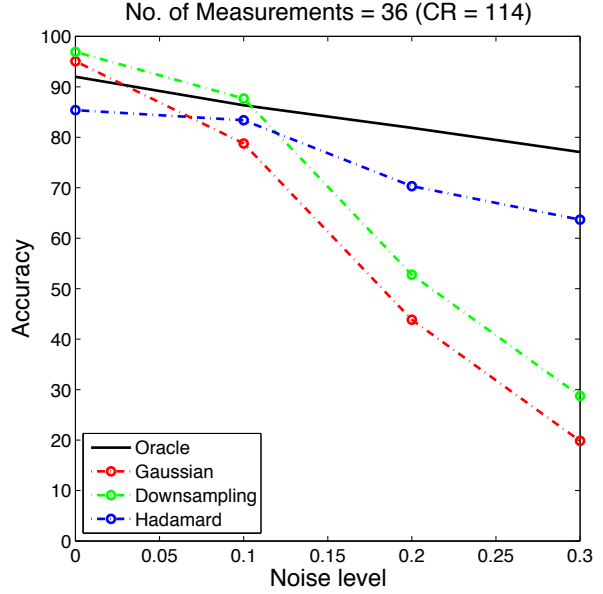
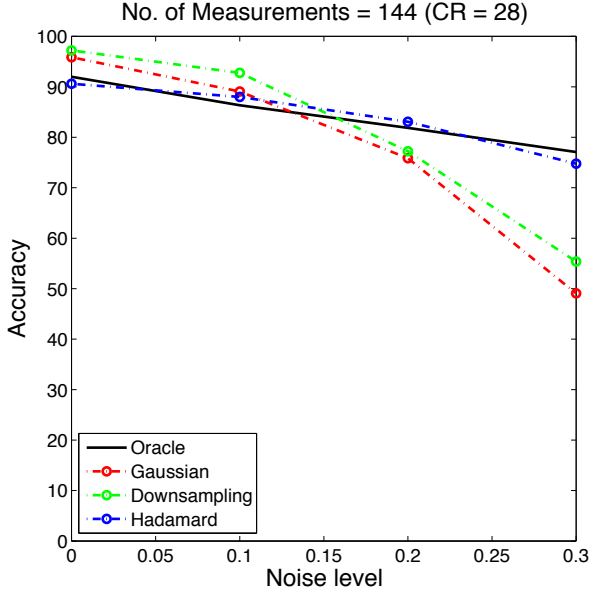


Figure 5. The figures show the variation of recognition accuracy for the AMP database for Oracle (no compression), Gaussian measurements, low-rank permuted Hadamard measurements, downsampling, for varying amounts of measurement noise. Note that results indicate that performance is close to Oracle for low-noise levels, and Hadamard is more stable in performance than Gaussian and downsampling operators.

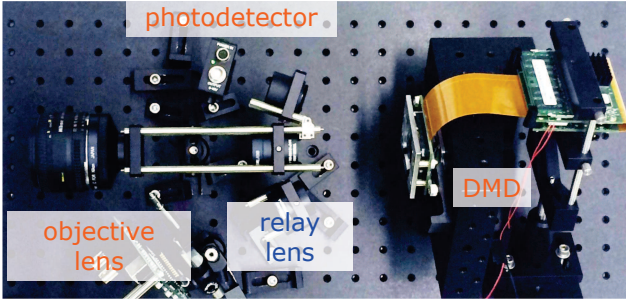


Figure 6. Photograph of the single-pixel camera used to obtain real data.

measurements in post-processing.

The new dataset consists of 120 face images, belonging to 30 subjects with 4 images per subject. For this, we printed out the images from a subset of the NIR database (used in the Section 3.1) and held each image in front of the SPC to obtain the compressive measurements. The images are captured using the SPC at a resolution of 128×128 . The dataset is divided into four train-test splits. For each split, the train set consisted of three images per subject, and the test set contained one image per subject. The recognition experiment was conducted at compression ratios (CR) of 100 and 200. The results are shown in Table 1.

Interestingly, face recognition accuracy is higher in the case of CRs of 100 and 200 than at a CR of 1. This is also the case with the controlled experiments with the NIR dataset (Figure 4). Similar phenomena have also been ob-

served in a related paper [24], as well as in general face recognition literature, where sometimes degradation of imagery leads to somewhat better recognition performance (perhaps due to suppression of high-frequency distractors). In our case, we believe that this effect may be attributable to the lower dimensional projection implicitly resulting in denoising, thus leading to better performance.

3.3. Reconstruction failure

Here, we demonstrate that, for high compression ratios (CR), inference is not possible even after reconstruction using state-of-the-art algorithms. Using the SPC measurements of a face image, we reconstruct the image using CoSaMP algorithm [21] at compression ratios of 5, 10 and 100 as shown in Figure 7. As inputs to the algorithm, we need to specify the measurement matrix, the sparsifying transform and the number of non-zero coefficients, s , required for the reconstructed image to be represented with high accuracy in the specified sparsifying transform. For all compression ratios, we use Daubechies wavelet transform as the sparsifying transform, and $s = \frac{M}{5}$, where M is the number of measurements. Clearly, reconstructed images at high CRs retain no valuable information that can be exploited for inference. Hence, we need to employ a framework – such as the one described in this paper – for direct inference on compressed measurements.

Compression ratio	No. of Measurements	Recognition Accuracy
1 (Oracle)	16384	60%
10	1638	62.5%
50	328	58.33%
100	164	53.33%
200	82	49.17%

Table 1. Face recognition results obtained on compressed measurements from a single-pixel camera. There is a graceful degradation in accuracy as CR is increased.

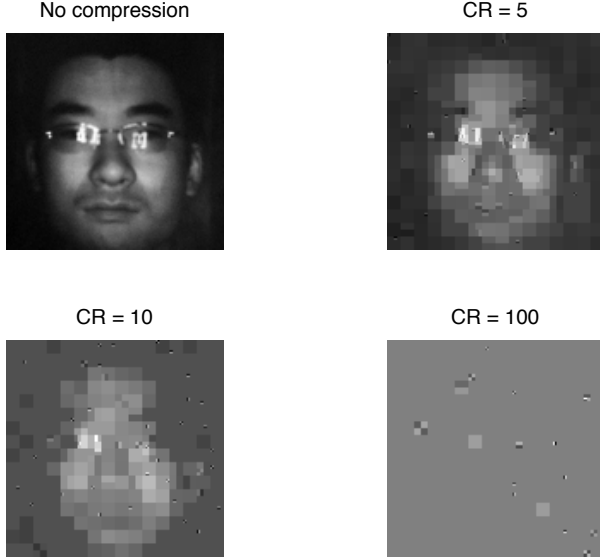


Figure 7. The figures show the reconstruction of images of a face at different compression ratios (CR) using the CoSaMP [21] algorithm. Note how reconstruction quality degrades very rapidly across compression rates which makes ‘reconstruction-then-inference’ a losing proposition.

4. Conclusions and future work

In this paper, we have addressed the problem of high-level visual inference from compressive cameras – which is a very specific example of the more general class of spatial-multiplexing imagers. In contrast to current interest in signal recovery, we ask whether visual inference is feasible without expensive signal reconstruction – the reconstruction often requiring high measurement rates and high signal to noise ratios. We show that indeed fairly robust inference is possible at compression rates and noise-levels where reconstruction fails miserably. We think this points to interesting research questions at the interface of computational imaging, computer vision, and machine learning. A few questions that warrant further research include: are reconstruction and inference only marginally related to each other? If so, what measurement operators might enhance inference capabilities as opposed to reconstruction capabilities?

References

- [1] T. Ahonen, A. Hadid, and M. Pietikäinen. Face recognition with local binary patterns. In *ECCV 2004*, pages 469–481. Springer, 2004.
- [2] P. K. Baheti and M. A. Neifeld. Adaptive feature-specific imaging: a face recognition example. *Applied optics*, 47(10):B21–B31, 2008.
- [3] R. Calderbank, S. Jafarpour, and R. Schapire. Compressed learning: Universal sparse dimensionality reduction and learning in the measurement domain. *Preprint 2009*.
- [4] E. J. Candes, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.
- [5] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.
- [6] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1):33–61, 1998.
- [7] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk. The smashed filter for compressive classification and target recognition. In *Electronic Imaging*, pages 64980H–64980H. International Society for Optics and Photonics, 2007.
- [8] O. Déniz, G. Bueno, J. Salido, and F. De la Torre. Face recognition using histograms of oriented gradients. *Pattern Recognition Letters*, 32(12):1598–1603, 2011.
- [9] D. L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.
- [10] K. Etemad and R. Chellappa. Discriminant analysis for recognition of human face images. *JOSA A*, 14(8):1724–1733, 1997.
- [11] W. B. Johnson and J. Lindenstrauss. Extensions of lipschitz mappings into a hilbert space. *Contemporary mathematics*, 26(189-206):1, 1984.
- [12] R. Juday. Application of correlation filters. In *Correlation Pattern Recognition*, pages 357–382. Cambridge University Press, 2005. Cambridge Books Online.
- [13] K. Kulkarni and P. Turaga. Reconstruction-free action inference from compressive imagers. *arXiv preprint arXiv:1501.04367*, 2015.
- [14] B. V. Kumar, A. Mahalanobis, and R. D. Juday. *Correlation pattern recognition*, volume 27. Cambridge University Press Cambridge, 2005.

- [15] S. Z. Li, R. Chu, S. Liao, and L. Zhang. Illumination invariant face recognition using near-infrared images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):627–639, 2007.
- [16] Y. Liang, W. Gong, Y. Pan, W. Li, and Z. Hu. Gabor features-based classification using svm for face recognition. In *Advances in Neural Networks–ISNN 2005*, pages 118–123. Springer, 2005.
- [17] A. Mahalanobis, B. Kumar, and D. Casasent. Minimum average correlation energy filters. *Applied Optics*, 26(17):3633–3640, 1987.
- [18] A. Mahalanobis and R. Muise. Object specific image reconstruction using a compressive sensing architecture for application in surveillance systems. *Aerospace and Electronic Systems, IEEE Transactions on*, 45(3):1167–1180, 2009.
- [19] A. Mahalanobis, B. Vijaya Kumar, S. Song, S. Sims, and J. Epperson. Unconstrained correlation filters. *Applied Optics*, 33(17):3751–3759, 1994.
- [20] P. Nagesh and B. Li. A compressive sensing approach for expression-invariant face recognition. In *IEEE CVPR*, pages 1518–1525. IEEE, 2009.
- [21] D. Needell and J. A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Commun. ACM*, 53(12):93–100, Dec. 2010.
- [22] M. A. Neifeld and P. Shankar. Feature-specific imaging. *Applied optics*, 42(17):3379–3389, 2003.
- [23] A. Rodriguez, V. N. Boddeti, B. V. Kumar, and A. Mahalanobis. Maximum margin correlation filter: A new approach for localization and classification. *IEEE Transactions on Image Processing*, 22(2):631–643, 2013.
- [24] A. C. Sankaranarayanan, P. K. Turaga, R. G. Baraniuk, and R. Chellappa. Compressive acquisition of dynamic scenes. In *ECCV 2010*, pages 129–142. Springer, 2010.
- [25] M. Savvides, B. V. Kumar, and P. Khosla. Face verification using correlation filters. *3rd IEEE Automatic Identification Advanced Technologies*, pages 56–61, 2002.
- [26] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *JOSA A*, 4(3):519–524, 1987.
- [27] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In *IEEE CVPR*, pages 1701–1708. IEEE, 2014.
- [28] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, 2007.
- [29] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk. An architecture for compressive imaging. In *IEEE International Conference on Image Processing*, pages 1273–1276. IEEE, 2006.
- [30] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.