Bulletin of Natural Sciences Research
Vol. 14, No. 1-2, 2024, pp. 50-59.

DOI: https://doi.org/10.5937/bnsr14-51559
Original Scientific Paper

APPLICATION OF COMPRESSIVE SENSING TECHNIQUES FOR ADVANCED IMAGE PROCESSING AND DIGITAL IMAGE TRANSMISSION

NENAD STEFANOVIĆ[1], BOBAN SAZDIĆ-JOTIĆ[2], VLADIMIR ORLIĆ[3], VLADIMIR MLADENOVIĆ[4], STEFAN ĆIRKOVIĆ[4]

[1]Center for Applied Mathematics and Electronics, Belgrade, Serbia
[2]Military Technical Institute, Belgrade, Serbia
[3]Vlatacom Institute, Belgrade, Serbia
[4]Faculty of Technical Sciences in Čačak, University of Kragujevac, Čačak, Serbia

ABSTRACT

The field of compressive sensing (CS) has emerged as a transformative approach in the acquisition and processing of high-dimensional data. This paper presents a comprehensive study on the application of compressive sensing techniques to advanced image processing and digital image transmission. By leveraging the inherent sparsity in natural images, CS allows for significant reductions in the amount of data required for accurate reconstruction, thereby overcoming the limitations imposed by the traditional Shannon-Nyquist sampling theorem. We explore the theoretical foundations of CS, including the principles of sparsity and incoherence, and provide a detailed overview of the Orthogonal Matching Pursuit (OMP) algorithm, a prominent greedy algorithm used for sparse signal recovery. Experimental results demonstrate the efficacy of CS in improving image reconstruction quality, as evidenced by enhancements in peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). Additionally, we discuss the practical implementation of CS in single-pixel cameras and its potential impact on future imaging technologies. The findings suggest that CS offers a robust framework for efficient image acquisition and processing, making it a valuable tool for various applications in multimedia, medical imaging, and remote sensing.

Keywords: Sensing, Reconstruction, Transmission, Sparsity, Algorithms.

INTRODUCTION

One of the most prevalent trends in contemporary technologies is the collection (acquisition), transmission, analysis, and processing of large amounts of data. Mass data processing is a common issue when it comes to multimedia data, medical and biomedical data, radar signals, and similar, and traditional sensor devices face strict requirements when performing acquisition according to the already standard Shannon-Nyquist criterion (sampling theorem), which states that the sampling rate of a signal must be at least twice the highest frequency present in the signal spectrum.

In the last decade, decade and a half, new methods and techniques known collectively as compressive sensing have been intensively developed, which surpass the boundaries of the existing sampling theory by applying the concept of data compression during the sampling/observing process itself. This reduces the amount of collected-measured data and thereby significantly reduces the engagement of available resources (time, energy, device complexity, etc.) on the end, i.e., the sensor side of the system.

This paper, in addition to describing the concept of compressive sensing, examines the application of a specific algorithm from the group of "greedy algorithms" for fast image reconstruction. The paper is divided into three main parts: theoretical framework with an intuitive approach, adaptation of the algorithm for advanced image processing, and presentation of the results achieved by applying such an algorithm on concrete examples and obtained performances from the perspective of peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM).

THEORETICAL PART

Compressive sensing (cs)

Definition and Basic Concept

The concept of compressive sensing has received different names over a relatively short period, depending on its application (compressive sampling, sub-sampling, sparse sensing, sub-space sampling, etc.), and a relatively small number of papers have been published in Serbian where the general terminology in this field has not been adopted. Taking into account all these applications and perspectives on the unique principle, the terms used in this paper are those the author considered appropriate translations for the given

*Corresponding author: stefan.cirkovic@ftn.kg.ac.rs

context. Thus, the term observation (of images) can in this case be considered synonymous with the term reading (of spectra) or sampling (of signals), while essentially retaining the same observational course.

The initial setup of the concept known as compressive sensing, presented in (Donoho, 2006), states that a signal can be reconstructed from only a small set of randomly measured values - samples, if the signal can be represented in a sparse (compressed) form in some transform domain. Sparsity, as a property of the signal, means that the signal can be represented in a certain (transform) domain with a small number of non-zero samples, implying fewer numerical computations, memory, and consumption for signal reconstruction. The second condition, of a technical nature, is the incoherence between the measurement (sampling) and the transformation (sparsifying) operator matrices.

The principle, initially adopted from image and signal compression techniques, can be compared with classical sampling theory and represented with a block diagram (Stankovic et al., 2016) in (Fig. 1):

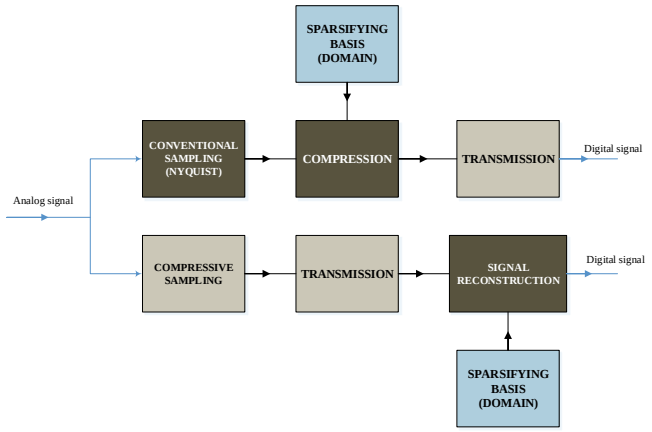

Figure 1. Comparative overview of classical and compressive sensing.

Parts of both systems are illustratively represented by blocks of different colors, with lighter shades representing blocks that are less burdened in terms of engaged resources. It is noticeable that in the case of compressive sensing, given the smaller amount of information collected, the focus of engaged resources shifts from the end points of the system (sensors, acquisition cards, high-resolution cameras, magnetic resonance scanners, geo-scanners, etc.) to the platform where signal reconstruction is performed.

The mathematical foundation that has led to a true "explosion" of publications and papers in the past decade lies in the premise that the process of data acquisition can be represented by a system of linear equations Eq. (1), (Foucart & Rauhut, 2013):

$$y = Ax \tag{1}$$

where from $y \in C^M$ the labeled vector of measured data (M - dimensional column), with $x \in C^N$ (N-dimensional column) vector of input data from the set of real numbers, and the matrix $A \in C^{M \times N}$ the operator of the linear measurement process, where $M<N$. The input signal is practically sampled at a smaller number of samples than the length of the input signal, resulting in an underdetermined system of linear equations (a greater number of unknowns – N than the number of equations – M). Assuming that the signal possesses the property of sparsity in the domain Ψ (which could be the domain of Fourier, DCT, Haar wavelet transform, etc.) and if the operator A has the corresponding properties, then it is possible to reconstruct the signal x from the vector y. For a more concise presentation of the theory, which can be found in numerous literature sources, we summarize the procedure step by step:

Sparsity of the signal: Let x represent an $N\times1$ column vector of the input analog signal discretized in time, as previously mentioned. For a given orthonormal basis matrix $\psi \in R^{N \times N}$, whose columns represent the basis elementary vectors $\{\psi_i\}_{i=1}^N$ x can be represented as a linear combination of basis vectors Eq. (2):

$$x = \sum_{i=1}^{N} \alpha_i \psi_i \tag{2}$$

or more compactly, $x=\Psi\alpha$, where α represents an $N\times1$ column vector of coefficients. These coefficients, in fact, represent the inner product of the vector x and the elementary basis vectors ψ_i, $\alpha_i < x$, $\psi_i > \psi_i^T x$, where $(^T)$ signifies the transpose operation. If the basis-domain Ψ mapping the input vector x to a vector with a small number, let's say K, of non-zero coefficients, then it can be written as

$$x = \sum_{i=1}^{K} \alpha_{n_i} \psi_{n_i} \tag{3}$$

the sparse version of the input vector, where ni are the indices of coefficients and basis vectors (referred to in literature as dictionary atoms Ψ) corresponding to K non-zero coefficients. Then α, a dimension $N\times1$, column vector with only K non-zero coefficients, denoted as $\|\alpha\|_0 = K$, where $\|\bullet\|_p$ denotes the lp-norm, or

$$\|x\|_p = \left(\sum_i \left| x_i^{\,p} \right| \right)^{\frac{1}{p}} \tag{4}$$

Calculating the norm of a specific vector or matrix represents different ways of measuring their length or magnitude (Patel et al., 2013). In the case where $p\to0$, the practically obtained l0-norm, namely the previously mentioned

number of non-zero coefficients of the vector, is obtained. Real signals cannot practically be considered sparse in any domain, but the vast majority possess compressive, compressible properties – meaning the property of being compressible or compressive signals. Compressibility implies that if the amplitudes of signal coefficients are arranged in a decreasing order $\alpha_{(1)} \geq \alpha_{(2)} \geq \ldots \geq \alpha_{(N)}$, results in their exponential decay

$$|\alpha|_n \leq C.n^{-s} \tag{5}$$

where $|\alpha|_{(n)}$ is the n^{th} largest coefficient value, $s \geq 1$, C constant. In that case, the error signal, or the difference between the signal obtained by a linear combination of basis vectors and the input signal, would also exponentially decay with the increase in the number of used coefficients

$$\|x_L - x\|_2 \leq CK^{-s\left(s-\frac{1}{2}\right)} \tag{6}$$

In other words, a small number of basic elementary vectors from Ψ can enable a precise approximation of x. Such approximation is known as nonlinear approximation.

Incoherence: The coefficients α_i from Eq. (2) are not directly measured in compressive sampling. Instead, the measurement of M ($M<<N$) projections of vector x is performed using a collection of vectors $\{\psi_i\}_{i=1}^{M}$, arranged as rows of an arbitrary permutation matrix Φ, dimension $M \times N$, resulting in a measured column vector y, dimension $M \times 1$, whose elements are $y_j = \langle x, \phi_j \rangle$. The measurement process can now be written as

$$y = \Phi x = \Phi \Psi \alpha = A\alpha \tag{7}$$

where A represents the measurement or observation matrix mentioned in Eq. (1). Reconstruction refers to the recovery of the sparse version of the input vector x, the vector α from which, by the inverse transformation Ψ^{-1}, an approximation of the input vector is obtained. An illustrative representation of the reconstruction, or obtaining the measured vector, is shown in (Fig. 2). In this case, the matrix Ψ is pictorially represented as a transformational matrix operator coefficients of the discrete cosine transformation *(DCT)*, commonly used for image processing and compression.

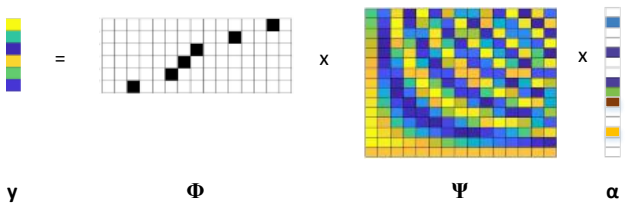

Figure 2. Illustrated representation of obtaining the measurement vector from compressively sampled signal.

For the algorithm to be successful, it is necessary for A to possess two fundamental properties:

The property of restricted isometry: known in the literature as *RIP (Restricted Isometry Property)*. For a matrix A to be said to possess the property of restricted isometry of order K, with constants $\delta_k \in (0,1)$, it must satisfy

$$(1-\delta_K)\|v\|_2^2 \leq \|Av\|_2^2 \leq (1+\delta_K)\|v\|_2^2 \tag{8}$$

for any vector v whose $\|v\|_0 \leq K$. An equivalent description would be that all subsets of K columns (vectors) taken from A are nearly orthogonal, implying that K-sparse vectors cannot lie in the null space of the matrix A. When *RIP* holds, A approximately preserves the Euclidean length (l_2-norm) of K-sparse vectors, i.e.,

$$(1-\delta_{2K})\|v_1 - v_2\|_2^2 \leq \|Av_1 - Av_2\|_2^2 \leq (1+\delta_{2K})\|v_1 - v_2\|_2^2 \tag{9}$$

applies to all K-sparse vectors v_1 and v_2. The related condition known as incoherence requires that the rows of Ψ cannot sparsely represent the columns of Ψ and vice versa.

Matrix incoherence: As a measure of mutual independence of constituent vectors for matrices Φ and Ψ is defined as

$$\mu(\Phi, \Psi) = \sqrt{N} \max \left| \langle \phi_i, \psi_j \rangle \right|, \quad 1 \leq i, j \leq N \tag{10}$$

The number μ represents the measure of similarity between two vectors of the matrix $A=\Phi\Psi$ and ranges between 1 and $\sqrt{N}$. The matrix A is said to be incoherent if μ is a very small number. Incoherence holds for various pairs of matrices, such as delta impulses and the Fourier basis, and with high probability between any matrix and a random matrix with Gaussian or Bernoulli distribution.

Reconstruction

As mentioned, algorithms for reconstructing compressively acquired signals involve finding a sparse approximation of the original input signal from compressive measurements, in an appropriate basis, framework, or dictionary. Research and development of various algorithms are motivated by reducing the number of measurements, noise resilience, speed, complexity, reliability, and credibility, etc. Algorithms are generally classified into six approaches for reconstruction, as shown in (Fig. 3), (Rani & Sushma, 2018).

The convex optimization approach involves finding the solution to formulation Eq. (1) using linear programming techniques. Some of the popular ones include the simplex algorithm, interior-point algorithm, Bregman algorithm, gradient projection for sparse representation (*GPSR*), fixed-point continuation, and others. These represent a global optimization approach.

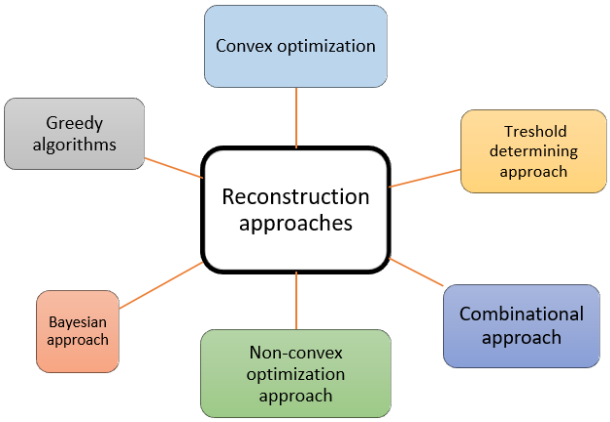

Figure 3. Approaches to compressive sensing reconstruction.

The approach using greedy algorithms, unlike convex ones, is an iterative step-by-step approach. During each iteration, the solution is updated by selecting only the columns of the reconstruction matrix that are highly correlated with the measurements. The selected columns are called atoms. Once selected, atoms are not used in subsequent algorithm steps, greatly reducing its computational complexity. The solution is obtained by iteratively choosing the best (closest to the original vector) solution in a "greedy" manner ("only the best is taken"), hence the name. These fall into the category of fast algorithms, but require some prior knowledge of the signal's sparsity measure. These algorithms can be further classified into two categories: serial and parallel.

The thresholding approach involves algorithms that work with K atoms from the reconstruction matrix, while also employing a certain thresholding sensitivity definition that sets all values below the threshold to zero, thereby reducing the impact of noise. The rest of the algorithm is very similar to the previously mentioned group. Some examples include Iterative Hard Thresholding, Iterative Soft Thresholding, Approximate Message Passing, etc.

Combinatorial approaches are primarily developed for finding sparse approximations during group testing, aiming to minimize the number of tests required to be conducted. They work on computing the minimum or mean values of measurements identified as constituent samples.

Non-convex optimization approach fundamentally differs from convex ones in the type of norm used to solve the minimization problem, specifically referring to norms located between l_0 and l_1. Examples of algorithms using this approach include solving focal undetermined system solutions and iteratively re-weighted least squares.

Bayesian approach is used for non-deterministic (stochastic) signals belonging to some probability distribution. Sparse Bayesian learning and others are known.

Orthogonal Matching Pursuit - OMP

Within the group of serial greedy algorithms, orthogonal matching pursuit (OMP) is included. Each iteration of these algorithms selects only one atom and calculates the corresponding non-zero element of the solution vector. The basic steps of these algorithms are depicted in the (Fig. 4), and the difference lies only in the solution update step:

Initialization: The residual vector r, dimension $M\times1$, together with the measurement vector y, are initialized to initial values. The solution, sparse vector α, dimension $N\times1$, and the index set Λ, dimension $M\times1$, are dimensioned as zero vectors, and the initial dictionary matrix D, which will be established during iterations, is initialized as a zero matrix. The cycle counter i is set to the value 1.

Atom search: In this step, the column of the reconstruction matrix A that has the maximum correlation with the residual vector r is found. The position of this atom is updated in the index set Λ.

Update of the sparse solution vector: Depending on the selected atoms, the solution set α_i is updated and the approximation of the measurement yi is found. The way the solution is updated is the step that distinguishes algorithms of this classification. In the OMP algorithm, this is the well-known least squares method.

Update of the residual: The new residual is calculated by subtracting the obtained approximation y_i (product of $D_i\alpha_i$) from the measurement vector y. These steps are repeated up to the desired sparsity level (K times) or until a desired residual value (less than some level $\varepsilon\geq 0$) is reached.

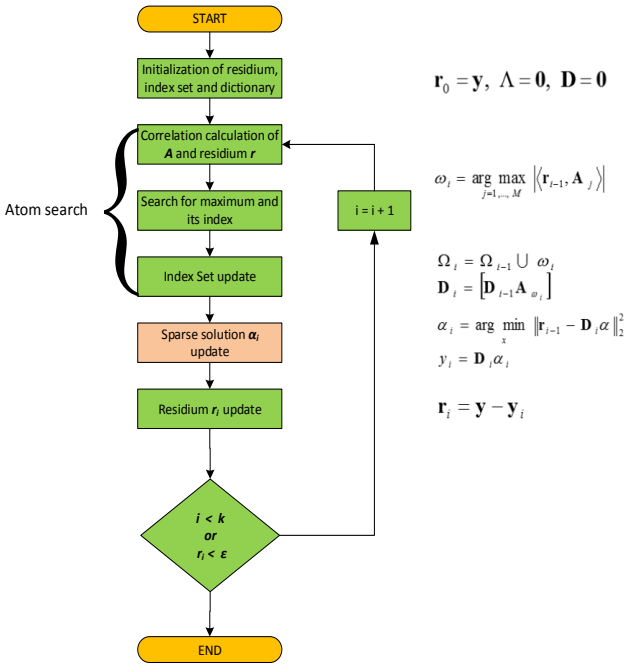

Figure 4. Algorithm for orthogonal matching pursuit.

For a clearer understanding of the principle of the orthogonal matching pursuit process, let's consider an example. The assumption is that for a given system

$$y = A\ \alpha$$

the measurement vector is known $y=\begin{bmatrix}1.65\\-0.25\end{bmatrix}$ and the matrix $A=\begin{bmatrix}-0.707 & 0.8 & 0\\ 0.707 & 0.6 & -1\end{bmatrix}$ and it is necessary to find α, which must be of dimension 3×1, and at the initialization phase, it is set as a zero-column vector $\alpha=\begin{bmatrix}0\\0\\0\end{bmatrix}$ whose coefficients are gradually updated during the steps.

We start from the premise that we consider the compressive matrix as a collection of basic column vectors or atoms:

$$A=\begin{bmatrix}-0.707 & 0.8 & 0\\ 0.707 & 0.6 & -1\end{bmatrix}=\begin{bmatrix}b_1 & b_2 & b_3\end{bmatrix}, \text{ where}$$

$$b_1=\begin{bmatrix}-0.707\\0.707\end{bmatrix}, b_2=\begin{bmatrix}0.8\\0.6\end{bmatrix}, b_3=\begin{bmatrix}0\\-1\end{bmatrix}.$$

If we denote that $\alpha=\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{bmatrix}$ then A is

$$\alpha=\begin{bmatrix}b_1 & b_2 & b_3\end{bmatrix}\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{bmatrix}=b_1\alpha_1+b_2\alpha_2+b_3\alpha_3=y$$

From this expression, it can be seen that the contribution of each atom from A depends on the values of the coefficient of the (un)sparse vector α. Thus, at the initial moment, it is necessary to find the atom that contributes the most to y. This process, for now, requires as many iterations as there are atoms.

1 step: The contribution of each atom is calculated by measuring the magnitude of its projection onto the measurement vector, i.e., by inner product.

$$\left|\langle b_1,y\rangle\right|=\left|\begin{bmatrix}-0.707\\0.707\end{bmatrix}\cdot\begin{bmatrix}1.65\\-0.25\end{bmatrix}\right|=$$

$$\left|-0.70\cdot(-0.25)+0.707\cdot1.65\right|=1.34$$

$$\left|\langle b_2,y\rangle\right|=\left|\begin{bmatrix}0.8\\0.6\end{bmatrix}\cdot\begin{bmatrix}1.65\\-0.25\end{bmatrix}\right|=1.17 \text{ and}$$

$$\left|\langle b_3,y\rangle\right|=\left|\begin{bmatrix}0\\-1\end{bmatrix}\cdot\begin{bmatrix}1.65\\-0.25\end{bmatrix}\right|=0.25$$

It can be seen that the largest contribution is from atom b_1, and therefore this vector is taken for updating the selected basis, whose coefficient is $\alpha=[\alpha_1]=-1.34$ (the coefficient of the most influential base vector in obtaining the final solution).

The contribution of each vector, of course, can be calculated in one step as $A^Ty=\begin{bmatrix}-1.34\\1.17\\0.25\end{bmatrix}$. The residual vector is calculated as $r=y-\alpha_1b_1=\begin{bmatrix}1.65\\-0.25\end{bmatrix}-(-1.34)\begin{bmatrix}-0.707\\0.707\end{bmatrix}=\begin{bmatrix}0.7\\0.7\end{bmatrix}$.

When the first vector is selected, the reconstruction matrix is updated and filled in a way that it adopts this vector as its new basis $D(1)=b_1=\begin{bmatrix}-0.707\\0.707\end{bmatrix}$.

2 step: Now, the calculation of contributions continues by searching for maximum correlations for the remaining atoms from matrix A (b_2 and b_3), but in relation to the residue vector r. In one step, we calculate contributions with $\begin{bmatrix}b_2 & b_3\end{bmatrix}^T\cdot r=\begin{bmatrix}0.8 & 0.6\\0 & -1\end{bmatrix}\begin{bmatrix}0.7\\0.7\end{bmatrix}=\begin{bmatrix}-0.98\\-0.7\end{bmatrix}$.

From this, it is concluded that the vector b_2 has the largest contribution in absolute value, based on which we choose this vector to add to the new base. In this second iteration, the new matrix is $D(2)=\begin{bmatrix}b_1 & b_2\end{bmatrix}$ and how close we are to the solution is calculated through least squares (Strang, 2016) $\min\left\|D(2)\cdot\alpha(2)-y\right\|_2$.

It is known that this problem in linear algebra can be solved with:

$$\alpha=D^+y,$$

where D^+ denotes the pseudo-inverse matrix of matrix D, for which $D^+=(D^TD)^{-1}D^T$. Now is $D^+(2)=\begin{bmatrix}-0.6062 & 0.8082\\0.7143 & 0.7143\end{bmatrix}$ (in MATLAB, the command is pinv), so at this moment (iteration) $\alpha=D^+y=\begin{bmatrix}-0.6062 & 0.8082\\0.7143 & 0.7143\end{bmatrix}\cdot\begin{bmatrix}1.65\\-0.25\end{bmatrix}=\begin{bmatrix}-1.2\\1\end{bmatrix}$.

The new residual will be

$$r=y-D(2)\alpha=y-\begin{bmatrix}b_1 & b_2\end{bmatrix}$$

$$\alpha=\begin{bmatrix}1.65\\-0.25\end{bmatrix}-\begin{bmatrix}-0.707 & 0.8\\0.707 & 0.6\end{bmatrix}\cdot\begin{bmatrix}-1.2\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Considering that we have reached orthogonal overlap (the residual between the measurement vector and the approximation is a zero vector), and knowing that the dimension of the vector is 3×1, we add as the last element 0, so $\alpha=\begin{bmatrix}-1.2\\1\\0\end{bmatrix}$.

When dealing with larger matrices, the process is iteratively repeated.

EXPERIMENTAL

Image processing using compressive sensing technique

Acquisition

One of the first devices that practically demonstrated the application of compressive sensing technique is the Rice single-pixel camera (SPC), (Patel et al., 2013). This camera essentially measures the inner product between an N-pixel sampled incident light beam from the scene and a set of N-pixel test functions, as shown in (Fig. 5).

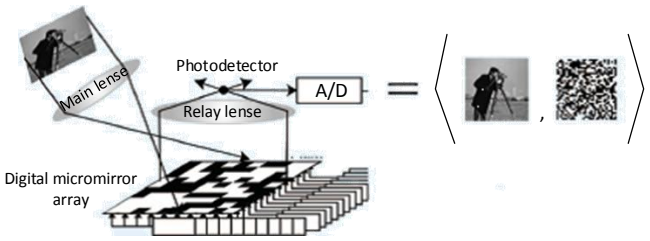

Figure 5. The operating principle of the Rice single-pixel camera, (Uzeler et al., 2013).

Such architecture utilizes only one photodetector to capture the scene. A digital micromirror device *(DMD)* array is used for the function of a pseudo-random binary sequence. The light beam is projected onto this array, and the beam intensity is measured with a photodetector. The orientations of the mirrors in the array can be rapidly changed, allowing for a series of different pseudo-random projections to be measured in a relatively short time. Reconstruction is then performed using some of the mentioned algorithms.

One of the main limitations of this architecture is that it requires the camera to focus on the object of interest until a sufficient number of measurements are collected, which can be restrictive for certain applications. Some of the other main architectures for compressive image sensing can be found in (Baraniuk et al., 2017).

Results of Reconstruction

For the quality of the reconstruction, the choice of the right transform domain in which the sparse image is observed is essential, i.e., the choice of the sensing matrix, as indicated by expressions Eq. (8) and Eq. (10). The Fourier domain is often used for the selection of the basis, especially for signals sparse in the frequency domain, followed by the Discrete Cosine Transform domain, Wavelet Transform domain (of various classes), etc. For the presentation of results in this work, two domains were comparatively used: Fourier *(F)* and Discrete Cosine *(DCT)*. Considering that one of the motivations for the work was the use of low-power cameras for terrain surveillance, an initial landscape image of dimensions 256×256 was used, followed by faces in the second part (known test image "Lena"), and a drawing image (parrot) of similar dimensions, as shown in (Fig. 6).

The platform used for simulation is based on an i7 processor with two cores at 2.8 GHz and 8 GB of RAM. The compressive sampling simulation is implemented by selecting the ratio of the measurement vector length to the sample size (M/N), i.e., how many pixel samples are taken for reconstruction. Initially, the Fourier matrix operator (identity, normalized) was chosen as the transformation operator. Variants with 60, 50, 40 percent of sampled pixels are shown in (Fig. 7) along with their corresponding reconstructions.

Figure 5. Original images used in the study: a) Lake b) Lena c) Parrot.

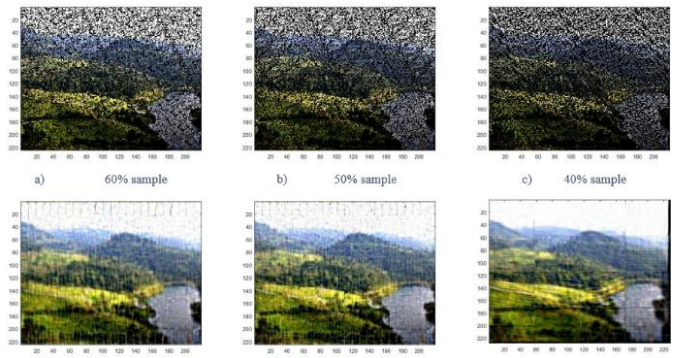

Figure 6. Reconstruction results by sparsifying in the Fourier domain with various M/N ratios.

For the sampled segment, column vectors of the two-dimensional image matrix were taken as samples, except in case (v) where, for clarity, the sampled vector was taken as a square block segment. In these cases, a relatively long reconstruction time was observed (several tens of minutes and over an hour for larger samples), primarily due to the large operator matrix and associated computational operations.

A better characterization of image sparsity was sought in the domain of discrete cosine transformation, considering that spectral analysis of one segment revealed the presence of a larger number of frequency components, as shown in (Fig. 8).

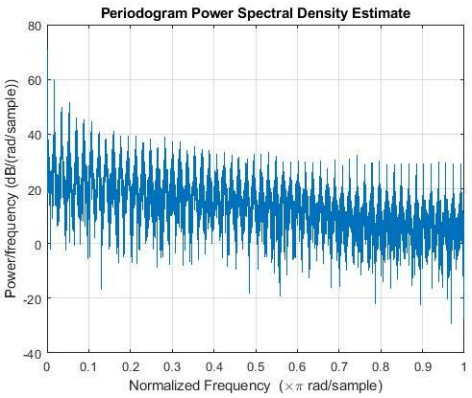

Figure 7. The spectral power density of one segment of the original image.

Even with just the replacement of the transformation matrix, a certain improvement was observed, as shown in (Fig.9).

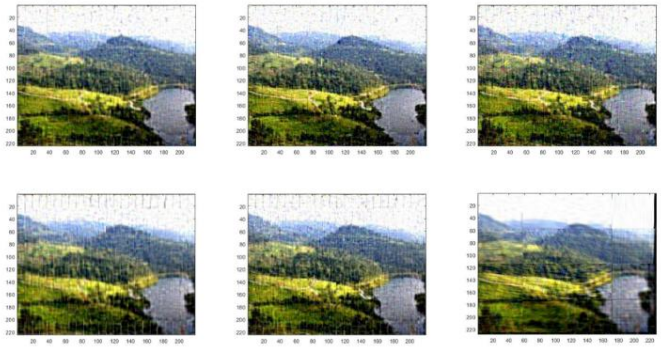

Figure 8. A comparative display of reconstruction with 60%, 50%, and 40% of points, in different sparse domains: top row DCT, bottom row FFT.

The reconstruction times, however, remained significant in terms of the required speed. For further improvements, it was necessary to enhance the reconstruction time, which initiated further exploration of possible solutions. One approach was to reduce the dimensions of the transformation matrix.

Algorithm Improvements

The proposed solutions to the problem, in terms of reducing the dimension of the transformation matrix, were investigated in (Gan, 2007) and (Sermwuthisarn & Parichat, 2009). In both cases, acquisition was considered block by block (dimension $n \times n$ pixels), but they differ in the approach to reconstruction. While in the first case, a Wiener filter, projection onto a convex set, and hard thresholding were used as reconstruction algorithms, in the second case, the approach of searching for orthogonal matches was used, called block-based OMP (Block-based Orthogonal Matching Pursuit). In this work, the latter method was used, only applied to color images and with some differences in parameterization.

When using block segments of the image, sized $n \times n$ pixels, the dimensions of the measurement matrix are expanded to $n_2 \times n_2$. For blocks of 64×64, the matrix is of size 4096×4096, for 16×16 blocks the matrix size is 256×256, while for 8×8 blocks the matrix size is 64×64.

Comparative reconstruction results by segmentation, with achieved reconstruction times, are shown in (Table 1), while the obtained appearances are shown in (Fig. 10).

Segmentation is implemented without overlapping between segments, but for further analysis, it can also be considered with overlap. Regarding the presented data in (Sermwuthisarn & Parichat, 2009), better times have been achieved, especially considering that in this work, the reconstruction of RGB color images compared to the monochromatic (grayscale) image used in the reference (dimensions are roughly similar), considering the triple pass through the algorithm (for each of the R, G, and B two-dimensional matrices individually). The table provides an approximate block size because to ensure proper segmentation, images are extended by a sufficient number of pixels divisible by the square root of the total number of segments.

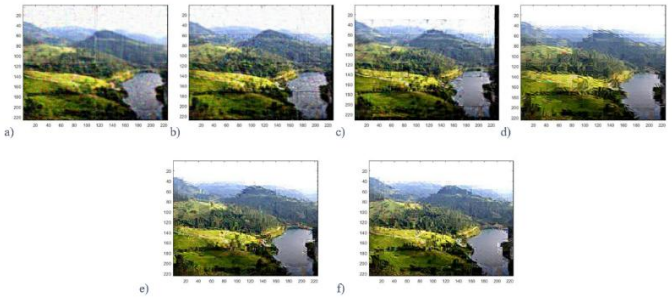

Figure 9. Reconstructions with 30% samples obtained by segmenting into segments: a) 4 b) 16 c) 64 d) 256 e) improvements of 256 segments, 3 times larger K f) improvements of 256 segments, 6 times larger K.

Table 1. Reconstruction time for different segment sizes and sparsity levels.

Point image	Block Size ($n \times n$)	M/N Ratio	Segments	Sparsity (K)	Reconstruction Time
a)	128×128	30	4	M/6	Several hours
b)	64×64	30	16	M/6	51.95 s
c)	32×32	30	64	M/6	10.58 s
d)	8×8	30	256	M/6	1.79 s

A better assessment of reconstruction quality is considered when it comes to face recognition. An example taken is the test image "Lena," which is often used in the references provided. Segmentation was performed on 256 blocks with different sparsity levels, (Fig. 11).

Figure 10. Image face reconstruction: in the first column with sparsity degree M/6, in the second M/2, in the third column M, the original image is framed on the right.

By observing, one can directly assess the recognizability of the face compared to the original (for example, the author has personally highlighted some of the reconstruction results with a red line). In step (Orthogonal Matching Pursuit – OMP) of the residue update, it has already been stated that the algorithm terminates after K iterations (sparsity level, thus directly affecting the reconstruction speed), or until the residual vector magnitude is reduced to some minimum match ($\varepsilon \geq 0$). The value of this minimum can be viewed as a number indicating the level of approximation and solution match, and for $\varepsilon=0$, the match is perfect.

In (Fig. 12), we notice that the level of matching actually determines the level of detail in the image, but also that details are not decisive, and sometimes a lower level of detail may be more useful for drawing conclusions about the necessary information from the image (whether it's a certain person or not, which person or characteristic it is). The reconstruction speed, logically, doubled in the case of $\varepsilon=550$ compared to the case with $\varepsilon=0$. When $\varepsilon=0$, we also notice small errors in the reconstruction (blue square at the far right and two yellow squares at the top and bottom). Errors appear in areas of the same color shades and when a complete level of matching is defined. The algorithm corrects itself by taking values greater thanzero.

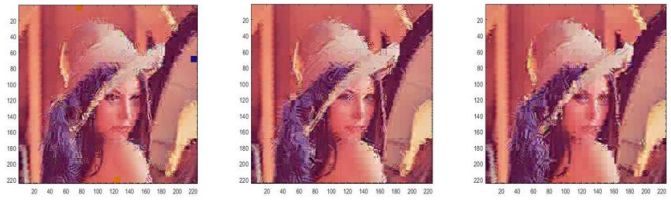

Figure 11. Reconstruction for different levels of matching, sequentially $\varepsilon = 0$ for 4.53s, $\varepsilon = 250$ for 2.45s, and $\varepsilon = 550$ for 2.19s.

The question of assessing the quality of reconstruction in the current analysis is based on subjective evaluation. The Peak-Signal-to-Noise-Ratio (PSNR) is one of the commonly used parameters for objectively measuring the quality of reconstruction in image codec (encoder/decoder) for image compression. PSNR is actually an approximation of the eye's perception of reconstruction quality. PSNR is most easily defined through the Mean Squared Error (MSE): for a given monochromatic image i and its noise-degraded approximation *K*, it is

$$SKG = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j)-K(i,j)\right]^2, \; PSNR \text{ (dB) is}$$

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{SKG}\right) = 20\log_{10}\left(\frac{MAX_I}{\sqrt{SKG}}\right) =$$

$$20\log_{10}\left(MAX_I\right) - 10\log_{10}\left(SKG\right)$$

where MAX_I is the maximum possible pixel value in the image. In our case, images with 8-bit pixel values were used, so the maximum value is 255. For color images with three RGB values per pixel, as used in the study, the same definition of the PSNR parameter is taken, except that MSE is calculated as the sum of sums for each color divided by the size of the image (dimensions $m \times n$) and multiplied by 3. A closer criterion to human eye perception is when this value is calculated on the luminance channel (Y channel of the YC_bC_r image format), as shown in (Fig. 13). The acceptable value of this parameter for lossy compression, according to the source (Thomos, et al., 2006), is from 20 to 25 dB.

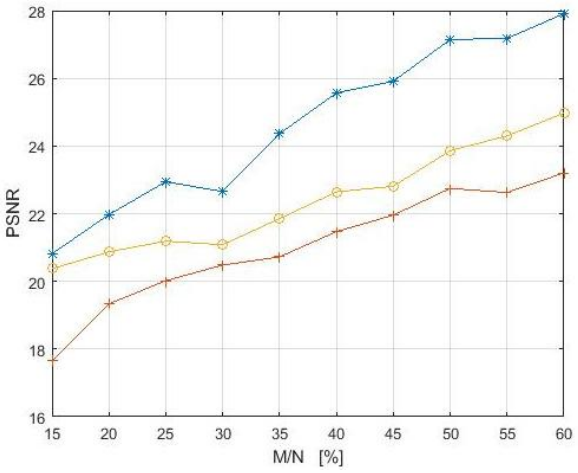

Figure 12. Peak Signal-to-Noise Ratio (PSNR) values for different types of images.

Another method considered here for quality assessment is measuring the Structural Similarity Index (SSIM) (Wang et al., 2004). This parameter describes the measure of similarity between two images, i.e., the perception of the resulting image change in its structure, and falls into the group of parameters for which complete referencing of the obtained reconstruction result with the original is necessary. Without delving into the details of the mathematical model, this parameter was used here as a complementary parameter to the traditional approach of measuring the peak signal-to-noise ratio. The concept of structural information in the image is related to the idea that pixels have strong mutual dependence when spatially close, and this dependency carries important information about the structure of objects in the visual scene (Thomos, et al., 2006). The value of the index ranges from -1 to 1 and is a decimal value, where 1 represents an identical set of data (perfect similarity), while 0 represents no structural similarity at all. The obtained index values for all three types of images used are shown in (Fig. 14).

In (Fig. 14), the change of Structural Similarity Index (SSIM) with the percentage of samples (M/N) is shown for three images: Lena, Parrot, and Nature. SSIM is a measure that describes how similar two images are in terms of their

structure, where a value of 1 indicates perfect similarity and a value of 0 means no structural similarity.

The X-axis represents the percentage of samples used for image reconstruction, ranging from 15% to 60%. The Y-axis shows the SSIM value, ranging from 0.4 to 0.85. Three lines on the graph represent three different images: Lena (marked with asterisks), Parrot (marked with plus signs), and Nature (marked with circles).

For the Lena image, SSIM starts around 0.55 at 15% of samples and gradually increases to approximately 0.85 at 60% of samples. This indicates that the reconstruction of the Lena image is very good, with a high level of similarity to the original. The Parrot image shows lower initial SSIM values, starting around 0.42 at 15% of samples and reaching about 0.7 at 60% of samples, indicating moderate similarity to the original image. The Nature image has an initial SSIM of around 0.46 at 15% of samples, gradually increasing to about 0.75 at 60% of samples, showing good similarity to the original but not as high as the Lena image.

Figure 14 shows that all SSIM indices increase with higher percentages of samples, indicating improved image reconstruction quality with increased sample usage. The Lena image exhibits the highest reconstruction quality, while the Parrot image is the most challenging to reconstruct. The Nature image falls somewhere in between in terms of reconstruction quality. These results demonstrate the effectiveness of the reconstruction algorithm used, as even at higher compression levels, the reconstructed images maintain acceptable levels of structural similarity to the originals.

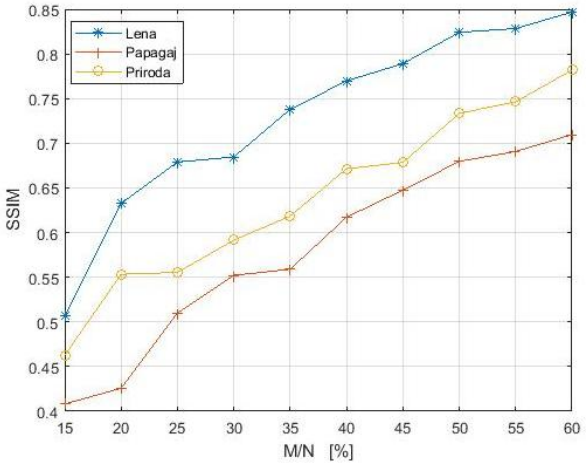

Figure 13. Values of the Structural Similarity Index for different types of images.

Based on the images, we can observe how the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) increase with the growth of sample percentage (M/N). From the presented results, it can be concluded that even under significant compression, the chosen algorithm maintains reconstructed approximations above the commonly acceptable threshold of 20dB PSNR. Additionally, the obtained SSIM values consistently align with previous findings (higher SSIM indicating a higher signal-to-noise ratio in the image). These parameters are utilized, for instance, in pattern recognition in images, faces, objects, etc.

CONCLUSION

The concept of compressive image sensing is presented in the paper, along with the basic conditions for its application, current approaches to reconstructing such signals, mathematical formulation with an intuitive example of orthogonal matching pursuit algorithm, and the method of image acquisition with compressive sampling. Furthermore, results obtained based on the applied algorithm are presented using three structurally different images: faces, natural landscapes, and drawings.

It can be concluded that acceptable results are achieved for a compression level of around 25-30% of samples (over 20 dB PSNR), and better structural similarity is obtained for faces compared to images of natural landscapes and drawings, which is important from the perspective of automated recognition. Input vectors have shown better sparsity properties in the Discrete Cosine Transform (DCT) domain compared to the Fourier domain (FFT). Some references claim that the compression level can be lowered to 4-10% using atoms from the Wavelet Transform dictionary, with a cautious note that the orthogonality of the matrix in this domain is not unequivocally determined and must be carefully considered during its design.

The method of forming the input vector also showed its impact on the algorithm's performance, primarily in terms of reconstruction time for the desired quality. Thus, the reconstruction results in this work are very comparable to the results in (Sermwuthisarn & Parichat, 2009). A wide range of possibilities remains open for the application of more advanced variants of this method, as well as improvements to the algorithm itself in some other sense. One of the possible improvements is presented in (Safavi, 2016). The relevance of the topic today is demonstrated by recent algorithm enhancements achieved (Zhao et al., 2019).

By employing compressive sensing/selecting methods, it is possible to form a signal at the end sensor points using simpler, cheaper, and low-power devices, which will enable satisfactory image reconstruction at the destination in the transmission system. This avoids the use of expensive hardware with high power consumption and provides an advantage over classical image compression methods. A small amount of data acquired at the end points lessens the burden on the transmission network and the connection of a large number of sensors to it.

ACKNOWLEDGMENTS

This study was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, and these results are parts of the Grant No. 451-03-66/ 2024-03/200132 with University of Kragujevac – Faculty of Technical Sciences Čačak.

REFERENCES

Baraniuk, R. G., Goldstein, T., Sankaranarayanan, A. C., Studer, C., Veeraraghavan, A. & Wakin, M. B. 2017. Compressive Video Sensing: Algorithms, architectures, and applications, IEEE Signal Processing Magazine, 34(1), pp. 52-66, Jan. 2017, doi: 10.1109/MSP.2016.2602099.

Donoho, D. L. 2006. Compressed Sensing. IEEE Transactions on information theory, 52(4), pp.1289-1306.

Gan, L. 2007. Block Compressed Sensing of Natural Images. In: Proceedings of the 2007 15th International Conference on Digital Signal Processing, Cardiff, UK, 1-4 July 2007.

Rani, M. & Sushma R. 2018. A Systematic Review of Compressive Sensing: Concepts, Implementations and Applications. IEEE Access, 6.

Patel, V.M., Vishal M. & Rajesh C. 2013. Sparse Representations and Compressive Sensing for Imaging and Vision. New York: Springer.

Foucart, S. & Rauhut, H., 2013. A Mathematical Introduction to Compressive Sensing. New York: Springer.

Sermwuthisarn, P., Auethavekiat, S., & Patanavijit. V. 2009. A fast image recovery using compressive sensing technique with block based orthogonal matching pursuit. In 2009 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS). IEEE. https://doi.org/10.1109/ispacs.2009.5383863

Stankovic, S., Orovic, I. & Sejdic, E. 2016. Multimedia Signals and Systems: Basic and Advanced Algorithms for Signal Processing. London: Springer.

Strang, G. 2016. Introduction to Linear Algebra, 5th ed. Wellesley: Wellesley - Cambridge Press.

Seyed, H. S., & Farah T-A. 2016. Sparsity-aware adaptive block-based compressive sensing. IET Signal Processing. pp. 36-42, https://doi.org/10.1049/iet-spr.2016.0176.

Thomos, N., Boulgouris, N. V., & Strintzis, M. G. 2006. Optimized transmission of JPEG2000 streams over wireless channels. IEEE Transactions on Image Processing, 15(1), 54-67. doi: 10.1109/TIP.2005.860338.

Uzeler, H., Cakir, S., & Ayta¸, T. 2013. Image generation for single detector infrared seekers via compressive sensing. Proceedings of SPIE - The International Society for Optical Engineering, 8896. doi: 10.1117/12.2030104.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. 2004. Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, 13(4), pp. 600-612, doi: 10.1109/TIP.2003.819861.

Zhao, L., Hu, Y., & Liu, Y. 2019. Stochastic Gradient Matching Pursuit Algorithm Based on Sparse Estimation. Electronics, 8(2), p165, doi: 10.3390/electronics8020165.