Random Projections and Their Applications in Computer Vision

Guilherme Schu, John Soldera, Rafael Sachett and Jacob Scharcanski

***Abstract*— Computer vision problems often require extracting and handling large volumes of high dimensional data. Commonly, a dimensionality reduction is initially applied to project the features representing the input data into lower dimensionality spaces before its analysis and/or classification. Techniques like Principal Component Analysis (PCA) have been used for dimensionality reduction, but data structure distortions may result from these projections, often leading to analysis and/or classification errors. On the other hand, random projections can provide dimensionality reduction while preserving the data local properties, and sample relative inter-distances in the lower dimensionality space. We review some of the theoretical foundations of random projection methods, and discuss their application in dimensional reduction and in computer vision problems. Also, some recent trends in the field of random projections are discussed.**

***Keywords*— random projections, computer vision, dimensionality reduction.**

I. Introduction

Computer vision problems often involve handling large amounts of high dimensionality data obtained by sensors like cameras or other devices. However, handling high dimensional data is computationally costly and can downgrade the performance of computer vision algorithms. Specially because as the data dimensionality increases, the data samples become more sparsely distributed in the feature space since the number of samples available remains the same. Therefore, dimensionality reduction techniques play an important role, since they can project the input high dimensionality data features to a lower dimensional space, while preserving important data structural characteristics.

Among several dimensionality reduction techniques available, Principal Component Analysis (PCA) is a method that is popular within the computer vision community. It allows to project high dimensionality data features into lower dimensional spaces relying on the possibility that most variability of the input data is restricted to a few directions in the original feature space. Nevertheless, if the initial space dimensionality is high (e.g, on the order of c*1000, with $c \geq 1$), and the final space dimensionality is much smaller (e.g. on the order of c'*1000, with $c' \ll 1$), PCA can be computationally expensive and may change substantially the local data structure [1].

Guilherme Schu is with Graduate Programme on Electrical Engineering, John Soldera formerly was with Graduate Programme on Computer Science and now is with Informatics Institute and Graduate Programme on Electrical Engineering, Rafael Sachett is with Graduate Programme on Computer Science, and Jacob Scharcanski is with Graduate Programme on Computer Science and Graduate Programme on Electrical Engineering, Federal University of Rio Grande do Sul, Porto Alegre, Brazil, E-mail: 'jacobs@inf.ufrgs.br'.
This work was partially supported by (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil).

The random projection (RP) approach is a different dimensionality reduction technique, which is becoming increasingly popular within the computer vision and pattern recognition community. RPs can obtain compact representations for high dimensionality data, preserving well the data sample inter-distances in the lower dimensionality space (i.e. are locality preserving transforms) at an accessible computational cost. In this paper, we outline the theoretical foundations of the RP approach in Section II. Also, some relevant applications of RPs in areas such as texture representation, biometrics and video processing are discussed in Section III. Finally, Section IV discusses some future research directions and presents our conclusions.

II. Theorectical Basics of Random Projections

Random projection (RP) is a projection technique where random projection matrices and linear matrix operations are used to project the data from a high dimensional space to a lower dimensional space [2], [3], [4]. Usually, RPs obtain compact representations for high dimensionality data, preserving the data sample inter-distances in the lower dimensionality space.

There are different ways to construct RP matrices. Different schemes could be used to construct a RP matrix $R \in \mathbb{R}^{d \times p}$ [4], and the values of its elements $R(i,j) = r_{ij}$ can be obtained as follows:

$$r_{ij} = \begin{cases} +1 & \text{with probability 1/2,} \\ -1 & \text{with probability 1/2.} \end{cases}$$

Recently, Medeiros *et al.* [5] proposed an alternative way to construct RP matrices, where each RP matrix element r_{ij} is obtained randomly as follows:

$$r_{ij} = \begin{cases} +1 & \text{with probability 1/2,} \\ 0 & \text{with probability 1/2.} \end{cases}$$

Given the $d \times p$ RP matrix R, the original $p \times 1$ data vector $\mathbf{x} \in \mathbb{R}^p$ is projected into the lower dimensionality space, obtaining $\mathbf{y} \in \mathbb{R}^d$ with $d < p$, by the following linear operation:

$$\mathbf{y} = R\mathbf{x}. \tag{1}$$

Therefore, RPs can be interpreted as a linear weighting of the original data vector components $\mathbf{x}$, where the weights r_{ij} are defined randomly [6]. More specifically, considering the original data vector $\mathbf{x} \in \mathbb{R}^p$ in the projection given in Equation 1, the random projection operation in Equation 1 can be expressed equivalently as :

$$\mathbf{y_i} = \sum_j \mathbf{r_{ij}} x_j, \tag{2}$$

where the elements of the j-th column of R are denoted $\mathbf{r_{ij}}$, the j-th component of $\mathbf{x}$ is denoted by x_j, and the random projection corresponding to the i-th row of R is denoted by $\mathbf{y_i}$ ($i = 1, \ldots, d$, and $j = 1, \ldots, p$). By increasing the number of rows in R, the dimensionality of the projected space is increased, and vice-versa. In Equation 1, each one of the d rows of R corresponds to the weights of a random projection, and the rows of R provide projections of the vector $\mathbf{x}$ in non-orthogonal projection directions (see also Equation 2).

Considering that data $\mathbf{x} \in \mathbb{R}^p$ lies in a $p - dimensional$ space, we should expect the number of non-orthogonal data projection directions $no \gg p$, where p is the number of orthogonal data projections directions. Nevertheless, the data representation based on a set of non-orthogonal random projections $\mathbf{y} \in \mathbb{R}^d$, where $d < p$, can approximate the representation of the original data $\mathbf{x}$ in a $d - dimensional$ space that has a lower dimensionality than the original space spanned by the set of orthogonal bases [6]. Besides, random projections lead to bounded errors in terms of the projected data positioning in the lower dimensionality space [3], since the data samples pairwise distances are preserved up to a factor $(1 \pm \epsilon)$:

$$(1-\epsilon)\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\epsilon)\|u-v\|^2, \tag{3}$$

where $f(u)$ is a function that maps the original data sample u to the lower dimensionality space and $f(v)$ is a function that maps the original data sample v to the lower dimensionality space. Therefore, given this positioning error guarantee in terms of data pairwise distances in the lower dimensionality space, the low dimensional space approximates the data structure in the original high dimensional space [7],[5].

An illustration of the RP mechanism applied to an image block is presented in Figure 1. A patch of size 3×3 is extracted from a grayscale image in Figure 1 a). The pixel values in the patch are sorted in ascending order obtaining the column vector $\mathbf{x} \in \mathbb{R}^p$ indicated by Figure 1 c). Figures 1 b) and d) show the process of generating a random matrix using the criterion in [5], and Figure 1 d) shows the $(d \times p)$ binary random projection matrix R. The random projection is obtained by multiplying R by the vector $\mathbf{x}$, according to Equation 1. The projected vector $\mathbf{y} \in \mathbb{R}^d$ is presented in Figure 1 e).

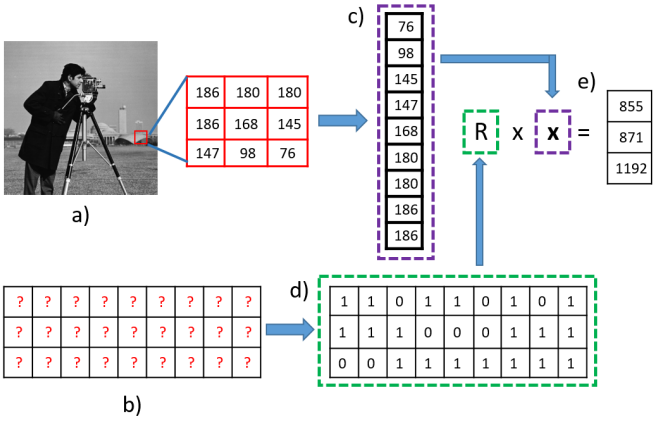

Fig. 1. Illustration of the random projection mechanism for an image block: a) sample image patch extracted from a grayscale image; b) random numbers generated using the criterion defined in [5], obtaining the random projection matrix R in d); c) 3×3 sampled patch is sorted in ascending order obtaining the column vector $\mathbf{x}$; e) projected data vector $\mathbf{y}$ computed by multiplying R by $\mathbf{x}$.

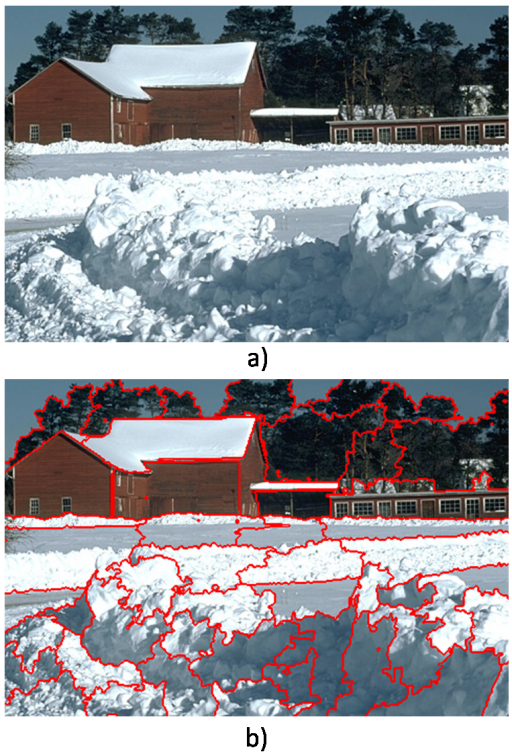

Fig. 2. Illustration of the segmentation of a natural scene using the method in [5], which is based on random projections of image patches: a) original image; and b) segmented image (region boundaries in red).

III. Applications of Random Projections in Computer Vision

The property of the RPs that data samples inter-distances are approximately preserved in a lower dimensionality subspace has been extensively explored [8]. Some RP applications in the field of computer vision, specially in biometrics, video processing and texture representation and segmentation, are briefly discussed next.

A. Biometrics

Biometrics-based verification systems confirm an individual identity based on the physiological and/or behavioral characteristics of the user [9]. Among various characteristics, biometric based methods need to account for two major issues: changeability and privacy. The changeability characteristic concerns the facility of users to use biometric representations for different applications, allowing biometric templates to

be reissued. Privacy, relates to the personal physiological characteristics of users. If one violates the storage device containing the data, the biometric templates should still preserve the personal information. Random projections have been used extensively for the generation of changeable and privacy preserving biometric templates.

In [10], the authors present an effective systematic analysis of RP as an intentional, repeatable, and non-invertible transformation for a changeable and privacy-preserving biometric template generation. The main idea is to reduce the dimensionality of biometric features while preserving privacy by sorting the random projected features. Recently, it has been proposed that iris recognition is performed based on an unified framework relying on RPs and sparse representations [11]. In order to guarantee the privacy of users physiological identity, the authors have followed the trend of incorporating RPs, and also random permutations, into their method to prevent the compromise of sensitive biometric information of the users.

In regard to face recognition, there are some challenging issues to be handled, specially in uncontrolled scenarios (e.g., airports, public areas, etc.), such as changes in illumination, pose and expression variations [12], [13],[14]. In order to account for such issues, face patterns representation often requires several dimensions, which raises a significant interest on RPs for dimensionality reduction [15], [16]. For instance, random transformations can be applied on the original features and the randomized vector can be sorted in descending order [17]. In this case, the authors argue that the sorting procedure makes it impossible to retrieve the original features, making the generated templates changeable and privacy-preserving.

B. Video Processing

The remarkable work made by Bingham and Mannila [18] yielded that projecting text and image data onto a random lower-dimensional space had recognition results comparable to conventional dimensionality reduction methods, such as PCA, since the RPs preserve similarity of data vectors. Such results, together with the mathematical framework given by [8], have been attracting researchers that work with high dimensionality spaces due to the easy computation of RPs as compared to other dimensionality reduction methods. Real time tracking problems also have been tackled using RPs [19], [20],[21]. For instance, it has been proposed in [22] that the restricted isometry property used in compressive sensing theory should be applied to obtain efficient projections of image features into a lower dimensionality compressed space for object tracking purposes. Samples are extracted from the frames and randomly projected into a lower dimensionality space, where the data is classified by a classifier that is learned online.

Random projections have also been explored in other video applications as well, such as anomaly detection [23], human action recognition [24], crowd density estimation [25], and video classification [26]. Particularly, it has been shown that RP methods can boost the efficiency of Orthogonal Matching Pursuit (OMP) algorithm [26]. The authors reveal a novel method that combines the work in [27] and the very sparse RP approach described in [4]. In video categorization problems, several interest points are extracted per frame, being each point described by feature vectors with many dimensions. The authors argue that, due the amount of data, efficient sparse projections algorithms play an important role in video classification [26].

C. Texture Representation and Segmentation

Texture features usually contain a composition of periodic and stochastic properties at various scales [28],[29]. Within this context, RPs can play an important role in texture feature representation. Recently, Medeiros *et al.* proposed a global image model texture descriptor based on a dictionary of stochastic features obtained by RP methods [30],[5]. Using this set of stochastic texture features along with a stochastic region merging algorithm, the authors showed that impressive image segmentation results can be obtained for unconstrained scenarios and medical applications (see Figure 2) [31],[32].

Texture and material classification using robust features by fusing multiple sorted RPs has been proposed in [33]. The authors propose to combine the texture features obtained using RPs as discussed in [34] and [35], and use support vector machine classifier (SVM) [36] to perform texture discrimination and classification. Also, RPs were proposed to project the image patches into a compressed space without loss of salient information in a global and information-preserving way [34]. Despite its simplicity, this method shows potential to outperform LBP and various filter bank-based methods. Liu *et al.* [35] proposed to reduce the rotation sensitivity of RPs by sorting random projections (SRP), which originated the concept of rotation-invariant stochastic feature textures. Combining the methods in [34] and in [35], even better texture classification rates were obtained for various challenging databases [33].

IV. Conclusions and Future Work

The interest on RP approach for dimensionality reduction problems has been increasing since its early development steps [3],[1]. This can be explained by the interesting properties of the RP approach, such as: (a) its accessible computation in comparison to other approaches; and (b) its capability of preserving approximate inter-distances of data samples in the lower dimensionality space obtained by RPs. Such properties have been attracting researches in the area of computer vision, from various fields, including biometrics and video processing, and also texture representation and segmentation [37].

The RP approach has an untapped potential to complement, or even replace, some of the methods currently used for feature extraction and representation in computer vision. Future work shall focus on new sparse image features obtained via RPs, that are more robust to local noise and artifacts. Such stochastic features are expected to help obtaining compact features spaces where computer vision tasks (e.g. biometrics, texture representation and classification, or even video processing) can be performed accurately, using reduced computational effort.

ACKNOWLEDGMENTS

The authors would like to thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil) for financial support.

REFERENCES

[1] H. Ritter and T. Kohonen, "Self-organizing semantic maps," *Biological cybernetics*, vol. 61, no. 4, pp. 241–254, 1989.

[2] N. Linial, E. London, and Y. Rabinovich, "The geometry of graphs and some of its algorithmic applications," *Combinatorica*, vol. 15, no. 2, pp. 215–245, 1995.

[3] W. B. Johnson and J. Lindenstrauss, "Extensions of lipschitz mappings into a hilbert space," *Contemporary mathematics*, vol. 26, pp. 189–206, 1984.

[4] D. Achlioptas, "Database-friendly random projections: Johnson-lindenstrauss with binary coins," *Journal of Computer and System Sciences*, vol. 66, no. 4, pp. 671–687, 2003.

[5] R. Medeiros, J. Scharcanski, and A. Wong, "Image segmentation via multi-scale stochastic regional texture appearance models," *Computer Vision and Image Understanding*, vol. 142, pp. 23–36, 2016.

[6] R. Hecht-Nielsen, "Context vectors: general purpose approximate meaning representations self-organized from raw data," *Computational intelligence: Imitating life*, pp. 43–56, 1994.

[7] L. J. Schulman, "Clustering for edge-cost minimization (extended abstract)," in *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*. New York, NY, USA: ACM, 2000, pp. 547–555.

[8] S. Dasgupta and A. Gupta, "An elementary proof of a theorem of johnson and lindenstrauss," *Random Structures & Algorithms*, vol. 22, no. 1, pp. 60–65, 2003.

[9] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on circuits and systems for video technology*, vol. 14, no. 1, pp. 4–20, 2004.

[10] Y. Wang and K. N. Plataniotis, "An analysis of random projection for changeable and privacy-preserving biometric verification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 5, pp. 1280–1293, 2010.

[11] J. K. Pillai, V. M. Patel, R. Chellappa, and N. K. Ratha, "Secure and robust iris recognition using random projections and sparse representations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pp. 1877–1893, 2011.

[12] J. Scharcanski, H. Proença, and E. Du, *Signal and image processing for biometrics*. Springer, 2014.

[13] J. Soldera, C. A. R. Behaine, and J. Scharcanski, "Customized orthogonal locality preserving projections with soft-margin maximization for face recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 9, pp. 2417–2426, 2015.

[14] C. A. Behaine and J. Scharcanski, "Enhancing the performance of active shape models in face recognition applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 8, pp. 2330–2333, 2012.

[15] N. Belghini, S. Ezghari, and A. Zahi, "3d face recognition using facial curves, sparse random projection and fuzzy similarity measure," in *2014 Third IEEE International Colloquium in Information Science and Technology (CIST)*. IEEE, 2014, pp. 317–322.

[16] J. Wang, "Efficient face recognition research based on random projection dimension reduction," in *2015 Seventh International Conference on Measuring Technology and Mechatronics Automation*. IEEE, 2015, pp. 415–418.

[17] Y. Wang and D. Hatzinakos, "On random transformations for changeable face verification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 3, pp. 840–854, 2011.

[18] E. Bingham and H. Mannila, "Random projection in dimensionality reduction: applications to image and text data," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2001, pp. 245–250.

[19] S. Zhang, H. Zhou, F. Jiang, and X. Li, "Robust visual tracking using structurally random projection and weighted least squares," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 11, pp. 1749–1760, 2015.

[20] A. Salaheldin, S. Maher, and M. Helw, "Robust real-time tracking with diverse ensembles and random projections," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2013, pp. 112–120.

[21] G. Tsagkatakis and A. Savakis, "Online distance metric learning for object tracking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 12, pp. 1810–1821, 2011.

[22] K. Zhang, L. Zhang, and M.-H. Yang, "Fast compressive tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 10, pp. 2002–2015, 2014.

[23] F. Chen, Z. Liu, and M.-t. Sun, "Anomaly detection by using random projection forest," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1210–1214.

[24] A. I. Maqueda, A. Ruano, C. R. del Blanco, P. Carballeira, F. Jaureguizar, and N. García, "Novel multi-feature bag-of-words descriptor via subspace random projection for efficient human-action recognition," in *Advanced Video and Signal Based Surveillance (AVSS), 2015 12th IEEE International Conference on*. IEEE, 2015, pp. 1–6.

[25] B. Xu and G. Qiu, "Crowd density estimation based on rich features and random projection forest," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2016, pp. 1–8.

[26] S. N. Vitaladevuni, P. Natarajan, R. Prasad, and P. Natarajan, "Efficient orthogonal matching pursuit using sparse random projections for scene and video classification," in *2011 International Conference on Computer Vision*. IEEE, 2011, pp. 2312–2319.

[27] N. Ailon and B. Chazelle, "Approximate nearest neighbors and the fast johnson-lindenstrauss transform," in *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*. ACM, 2006, pp. 557–563.

[28] J. Scharcanski and C. Dodson, "Texture analysis for estimating spatial variability and anisotropy in planar stochastic structures," *Optical Engineering*, vol. 35, no. 8, pp. 2302–2309, 1996.

[29] J. Scharcanski, "Stochastic texture analysis for monitoring stochastic processes in industry," *Pattern Recognition Letters*, vol. 26, no. 11, pp. 1701–1709, 2005.

[30] R. Medeiros, J. Scharcanski, and A. Wong, "Natural scene segmentation based on a stochastic texture region merging approach," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 1464–1467.

[31] A. Wong, J. Scharcanski, and P. Fieguth, "Automatic skin lesion segmentation via iterative stochastic region merging," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 6, pp. 929–936, 2011.

[32] J. Scharcanski and M. Celebi, *Computer vision techniques for the diagnosis of skin cancer*. Springer, 2014.

[33] L. Liu, P. W. Fieguth, D. Hu, Y. Wei, and G. Kuang, "Fusing sorted random projections for robust texture and material classification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 3, pp. 482–496, 2015.

[34] L. Liu and P. Fieguth, "Texture classification from random features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 3, pp. 574–586, 2012.

[35] L. Liu, P. Fieguth, D. Clausi, and G. Kuang, "Sorted random projections for robust rotation-invariant texture classification," *Pattern Recognition*, vol. 45, no. 6, pp. 2405–2418, 2012.

[36] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[37] J. Scharcanski, "Bringing vision based measurements into our daily life: A grand challenge for computer vision systems," *Frontiers in ICT*, vol. 3, p. 3, 2016.