Suitability of data representation domains in expressing human motion radar signals

Branka Jokanović, *Student Member, IEEE,* and Moeness Amin, *Fellow, IEEE*

***Abstract*—Human motion recognition (HMR) plays a key role in various fields including health monitoring. Radar is a type of sensor that has shown remarkable success in the classification of human motions. Different data representation domains have been used for the analysis of radar returns. Each domain provides one aspect of the observed motion not readily discernible in other domains. In this letter, we propose an approach to quantify the domain suitability in representing a human motion. In order to evaluate the proposed approach, we consider the time-frequency domain and the range map. Additionally, based on the demonstrated domain effectiveness, we propose a classification scheme which seeks to incorporate each domain consistent with its offerings. Experimental results show the importance of investigating domain offerings prior to the classification process.**

***Keywords*—Classification, Doppler, human motion recognition, principal component, range map, time-frequency domain**

I. Introduction

Much research in recent years has been focused on the use of human motion recognition (HMR) in health monitoring [1]-[3]. Short-term health monitoring is typical for rehabilitation services, while long-term monitoring is common in residential care facilities for the elderly.

Remote sensing of human motions offers several advantages over wearable devices, most notably, it is user independent. Camera-based systems capture small movements, but preprocessing methods, like background subtraction and extraction of human target, have to be employed before motion analysis. Privacy concern issue, associated with camera-based systems, can somewhat be alleviated by the use of systems where only the contours of human shape can be seen, e.g. infrared monitors. However, other issues like improper illumination and occlusions remain challenging.

Radar systems provide contactless monitoring that is not susceptible to lighting conditions or typical indoor occlusions [4], [5]. It is a proven technology for monitoring different targets, including humans and animals. For many years, radar has been used for target detection and tracking, and more recently, classification [6]-[12].

Authors are with the Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA e-mail: (branka.jokanovic, moeness.amin@villanova.edu).

This paper is made possible by NPRP Grant # NPRP 6-680-2-282 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

The time-frequency (TF) signal representation of radar returns is commonly used for observing human motions [7]. Features extracted in the TF domain often have physical interpretations, including energy, periodicity, and highest frequencies. Other domains have also been considered for the same purpose [13], [14]. For example, cadence velocity diagram can be used to detect abnormal gait [15] and extract features based on pseudo-Zernike moments [13].

Different motions manifest themselves differently depending on the domain used. Since there is no one single domain that provides distinctive signatures for all motion types, an effective HMR approach should undertake multiple data domain representations. Deciding on a proper domain for a given motion would depend on the resemblance/distinction between the respective motion signature and other motion signatures in the same domain. In this letter, we propose an approach that measures a motion distinction level in each domain with respect to other motions. This approach can be used prior to any classification process to guide domain selection and fusion schemes aiming at minimizing the classification errors intrinsic to a single domain.

The letter is organized as follows. Section II describes the dataset and the employed representation domains. The proposed approach for measuring motion distinction level is discussed in Section III. Experimental results are shown in Section IV, while the conclusion is given in Section V.

II. Domains used for radar return signals analysis

When using FMCW radar, it is possible to represent the radar signals in different joint-variable domains that include slow-time, range, and Doppler [14]. Each domain provides information that may not be present or easily identifiable in other domains. In this letter, we observe the data in the TF domain and the range map. These two domains are commonly used to represent radar returns and they provide valuable information about the target over time. The TF domain depicts the target velocity change, in our case the human torso and limbs, over time, whereas the range map provides information about the motion range extent and its time-dependence.

The spectrogram is used as the TF signal representation, depicting the distribution of the signal power over time n and frequency k. The spectrogram of a periodic version of a discrete signal $s(n), n = 0..N-1$ is given by,

$$SPEC(n,k) = |\sum_{m=0}^{N-1} h(m)s(n-m)e^{-j2\pi km/N}|^2, \quad (1)$$

where $h(m)$ is a window function.

In this work, we observe four common human motions: falling, sitting, bending and walking. Each motion is observed during a time span of 4s. Fig. 1 depicts four observed motions in the two aforementioned domains. It is clear that motions which can be visually distinguished in one domain may not be easily identified in another domain. For example, falling and sitting can be confused in the TF domain, while they can be readily distinguished in the range map.

III. Analysis of human motions in a chosen domain

In this section, an approach to quantify the distinction between a pair of motions is proposed. We find the principal components of each motion class. Similarity between motion classes then amounts to measuring closeness between the corresponding subspaces.

A. Measuring motion distinction level

Define matrix X that contains raster scanned images $x_i, i = 1, \ldots, M$ of a specific motion,

$$X = \left(\begin{array}{c|c|c|c} x_1 & x_2 & ... & x_M \end{array} \right). \tag{2}$$

We assume that X is normalized. Each vectorized image is of size 1xN. The principal components are defined by the eigenvectors of the covariance matrix $C_{XX} = XX^T$. Various methods are used to estimate the covariance matrix and to compute the principal components [16], [17]. The maximization of the variance is one of the well-known approaches where the strongest principal eigenvector ν is obtained by solving

$$\begin{aligned} \max_{\nu} \quad & \nu^T C_{XX} \nu \\ \text{subject to} \quad & \|\nu\|^2 = 1. \end{aligned} \tag{3}$$

Other components can be obtained in an iterative manner, by removing the already computed principal components from the signal and applying the same procedure until all components are extracted. The number of components that can be used for quantifying the motion level distinction depends on the importance of principal components. Fig. 2 shows the three strongest principal components of the observed motions in image form. Finding a class subspace is an important step in the classification process. The aim here is to quantitatively describe the suitability of each domain for a given motion. Towards this end, we measure similarity between the motion subspaces using both Canonical correlations and Pearson correlation coefficients.

B. Similarity measures

Canonical correlations are a well-known tool for measuring the distance between two subspaces [18], [19]. Define $\mathcal{U}$ and $\mathcal{V}$ as d-dimensional subspaces belonging to two motion classes. Canonical correlations are cosines of canonical angles $\theta_i, i = [1, ..., d]$ (also known as principal angles), and are defined as:

$$\begin{aligned} cos(\theta_i) = \max_{u \in \mathcal{U}, v \in \mathcal{V}} \max \quad & |u_i^T v_i| \\ \text{subject to} \quad & \| u \| = \| v \| = 1 \\ & u_i^T u_j = v_i^T v_j = 0, i \neq j. \end{aligned} \tag{4}$$

The solution of this problem can be obtained using Singular Value Decomposition (SVD) of matrices U and V that form orthonormal bases for the subspaces $\mathcal{U}$ and $\mathcal{V}$, i.e.,

$$U^H V = P \Lambda Q. \tag{5}$$

The columns of P and Q are the left- and right-singular vectors, respectively, whereas Λ contains singular values. These singular values are canonical correlations, i.e., $cos(\theta_i) = \lambda_i, i = 1, ..., d$. The minimum canonical angle is used to measure the closeness of two subspaces.

Another measure of similarity between two subspaces is to treat their respective components collectively instead of separately, as in the canonical correlations. That is, for each subspace, the corresponding d components of length N are averaged. In this way, two subspaces are represented by vectors x and y, respectively. The similarity can then be measured by finding the Pearson correlation value, defined as [20],

$$\rho = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \overline{y})^2}}, \tag{6}$$

where $\overline{x}$ and $\overline{y}$ are the sample means.

Figs. 3 (a,c) show both canonical and Pearson correlation coefficients in different domains when using the strongest principal component of each motion class. Figs. 3 (b,d) show the correlation coefficients corresponding to the five dominant principal components for spectrograms and range map, respectively. We can notice that in both cases, the high correlation coefficients properly describe the similar motions, i.e., falling and sitting in the TF domain, and bending and sitting in the range map. Conversely, and accordingly to the two measures, walk and bending motions are considered most distinctive in both representation domains. By increasing the number of principal components, the motion distinction level becomes more accurate. For example, the Pearson correlation coefficient better captures the similarity between falling and walking in the range map.

For comparison, we compute the Pearson correlation coefficients for the spectrograms, depicted in Fig. 1, in lieu of their subspaces. The results are shown in Fig. 4. It is evident that attempting to measure motion resemblance, based on the respective spectrograms, fails to provide a desired motion distinction level.

The results of Fig. 3 highlight the importance of representing human motions in the proper domains. They demonstrate two effective measures of motion distinction which can guide the selection of data representation prior to any classification process.

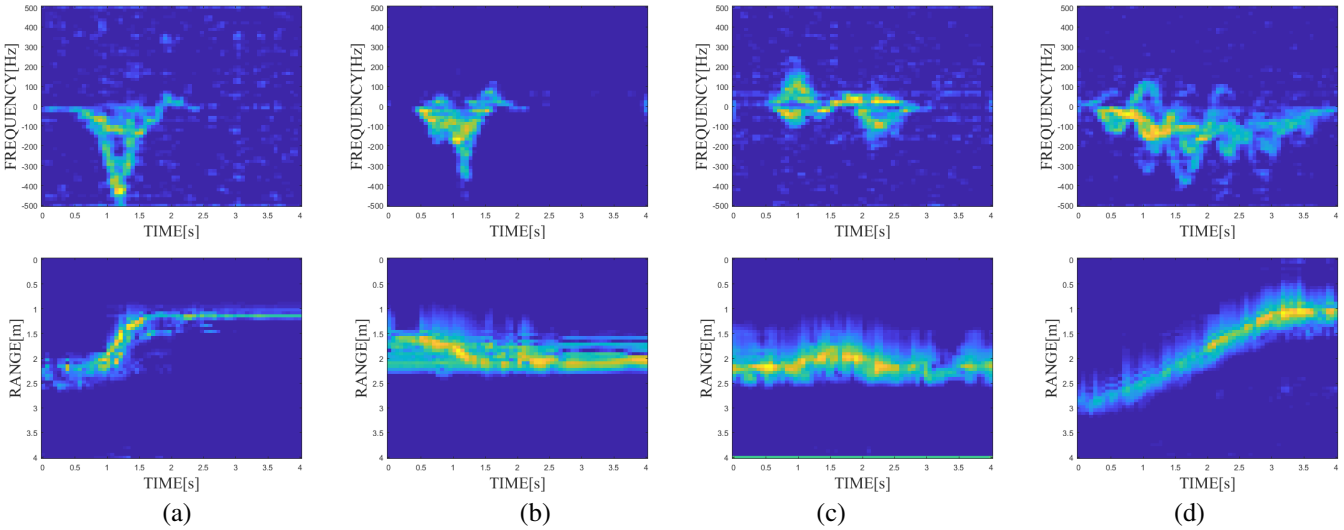

Fig. 1. Human motions in the TF domain (first row) and the range map (second row): (a) Fall, (b) Sit, (c) Bend, (d) Walk.

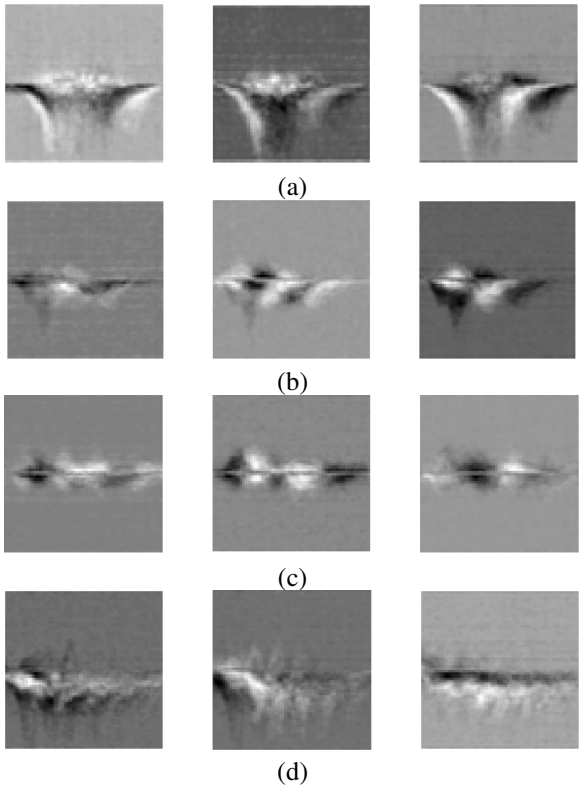

Fig. 2. Three strongest principal components of the observed human motions: (a) falling, (b) sitting, (c) bending and (d) walking.

C. Domain combination scheme for motion detection

The above analysis allows us to devise more appropriate domain combination schemes without relying on the visual assessments. In essence, it is prudent to group similar motions and regard them as one, instead of trying to detect each motion in every domain separately. Based on the results obtained in the previous section, proper motion groups are:

- Time-frequency domain: fall-sit;
- Range map: fall-walk, sit-bend.

One has the option to observe these domain representations in a parallel or serial manner. In this work, a serial configuration is proposed (Fig. 5). The test motion is first classified based on the range maps into two classes: fall-walk and sit-bend. This classification can be described as distinguishing between in-place motions versus motions that cause range translation. The next step is more precise classification, i.e., based on the radar return TF representation, motion is declared as either fall or walk (sit or bend).

It should be noted that this approach can be used in combination with any classifier or feature extraction method. In this work, we use Principal Component Analysis (PCA) based classifier to discriminate motions according to the scheme in Fig. 5 [21], [22]. PCA based classification consists of two stages, the training stage followed by the testing stage. The classification process can be described by the following steps:

Training

1: Vectorize each input image and stack them as columns in a matrix S
2: Normalize the matrix S by subtracting the average
3: Perform the eigendecomposition of the covariance matrix SS^T and select dominant eigenvectors
4: Project the training images onto the subspace spanned by selected eigenvectors

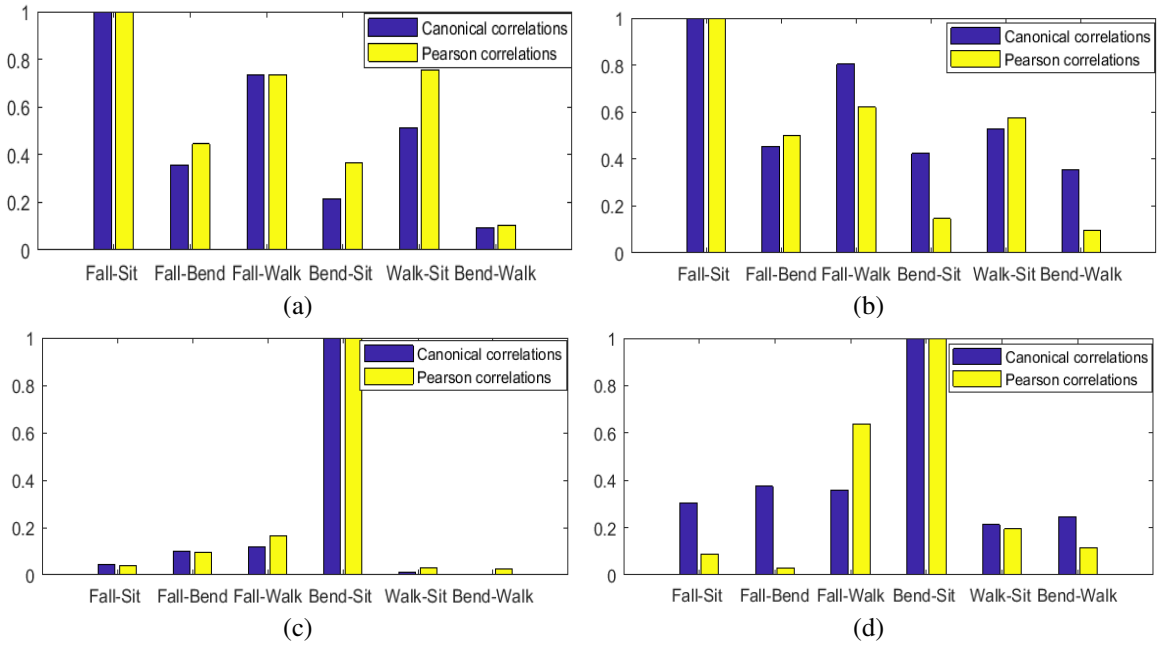

Fig. 3. Canonical and Pearson correlation coefficients for a pair of subspaces when 1 (a,c) and 5 principal components (b,d) are used. First row - spectrograms, second row - range maps. Correlation coefficient values are normalized in the range [0,1].

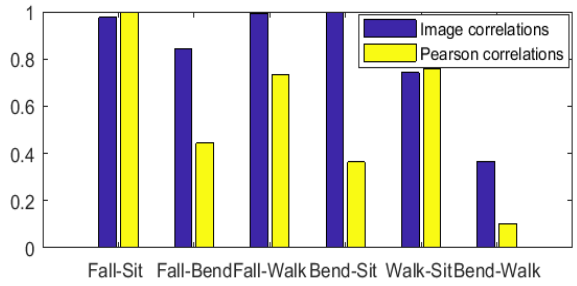

Fig. 4. Correlation coefficient for spectrograms. Pearson coefficients are included for the reference.

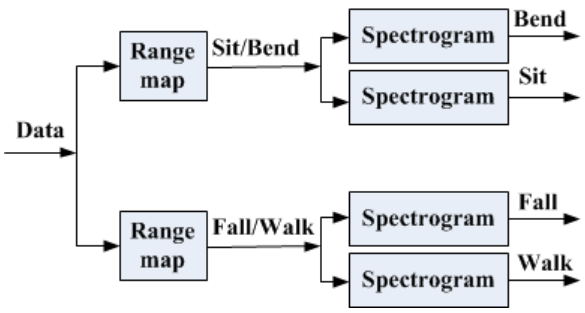

Fig. 5. Proposed combination scheme. The test data is first classified based on the range extent, followed by the classification based on the TF representation.

Testing

1: Project the testing image onto the subspace obtained in the training process
2: Compare the testing image projection with all projections obtained for the training images
3: The minimum Euclidean distance determines the class with the closest match to the observed test motion

IV. Experimental Results

The FMCW radar experiments were conducted in the Radar Imaging Lab at the Center for Advanced Communications, Villanova University. The radar system used in the experiments, named SDRKIT 2500B, is developed by Ancortek, Inc. The center frequency is 25 GHz, whereas the bandwidth is 2 GHz which provides 0.075 m range resolution. The obtained dataset is analyzed in Matlab 9.0.

The dataset contains 260 samples, 65 samples for each motion class. For the PCA based classifications, the training set consists of 50 samples for each class (200 samples in total), and each sample is a 64x64 image. Thus, for each motion class, 4096x50 matrix X is generated according to Eq. (2). This training dataset is also used to generate results in Figs. 3 and Fig. 4. The rest of the dataset is used for testing. The number of principal components was determined based on the eigenvalues. The success rate can be improved by including additional principal components [23]. However, after a certain number of components, there is no significant increment in the success rate. Based on the eigenvalues for the observed dataset, ten dominant principal components are used for the TF representation, whereas five dominant components are used for the range maps. When only using five components for the TF representation, a success rate of 73.3% is achieved, whereas the use of fifteen components yields negligible improvement (84.75%).

The confusion matrices for spectrograms and range maps are shown in Tables I and II, respectively. These tables show the correspondence between the actual and the classified class. For example, Table II depicts that the 93% of actual falls are classified as falls, while the rest (7%) are declared as

walk. These results underline the need to perform domain analysis prior to classification. It can be observed that the low success rate for the range maps is primarily caused by sitting and bending misclassification. Table III contains results when the proposed combination scheme is used. We can notice that domain combinations, when performed properly, improve the overall success rate. Additionally, the difference in computational times between the use of a single domain and the use of a combination scheme is not significant.

TABLE I. CONFUSION MATRIX FOR THE TF DOMAIN. SUCCESS RATE IS 84.5%, COMPUTATIONAL TIME IS 0.8 s.

Classified/Actual Class	Fall	Sit	Bend	Walk
Fall	86%	20%	-	7%
Sit	14%	80%	14%	7%
Bend	-	-	86%	-
Walk	-	-	-	86%

TABLE II. CONFUSION MATRIX FOR THE RANGE MAP. SUCCESS RATE IS 81.25%, COMPUTATIONAL TIME IS 0.5 s.

Classified/Actual Class	Fall	Sit	Bend	Walk
Fall	93%	7%	-	-
Sit	-	66%	34%	-
Bend	-	27%	66%	-
Walk	7%	-	-	100%

TABLE III. CONFUSION MATRIX FOR THE APPROACH BASED ON THE TREE-LIKE STRUCTURE. SUCCESS RATE IS 91.25%, COMPUTATIONAL TIME IS 1.4 s.

Classified/Actual Class	Fall	Sit	Bend	Walk
Fall	93%	-	-	14%
Sit	-	100%	14%	-
Bend	-	-	86%	-
Walk	7%	-	-	86%

V. CONCLUSION

In this paper, we introduced an approach which measures domain representation appropriateness for a given human motion. This approach provides an effective means to guide selection and combinations of multiple domain representations in order to minimize errors inherent in a single domain-based classifications. The proposed approach is evaluated using the TF domain and the range map. Based on the analysis of the observed domains, an effective serial combination scheme is proposed. Results demonstrate that proper combination schemes outperform the use of single domains.

REFERENCES

[1] A. I. Cuesta-Vargas, A. Galán-Mercant, and J. M. Williams, "The use of inertial sensors system for human motion analysis," *Phys. Ther. Rev.*, vol. 15, no. 6, pp. 462–473, 2010.

[2] C.-C. Yang and Y.-L. Hsu, "A review of accelerometry-based wearable motion detectors for physical activity monitoring," *Sensors*, vol. 10, no. 8, pp. 7772–7788, 2010.

[3] M. G. Amin, Y. D. Zhang, F. Ahmad, and K. D. Ho, "Radar signal processing for elderly fall detection: The future for in-home monitoring," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 71–80, 2016.

[4] F. Fioranelli, M. Ritchie, and H. Griffiths, "Classification of unarmed/armed personnel using the NetRAD multistatic radar for micro-Doppler and singular value decomposition features," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 9, pp. 1933–1937, 2015.

[5] A. Bouzerdoum, F. H. C. Tivive, and J. Fei, "Through-the-wall radar signal classification using discriminative dictionary learning," in *IEEE ICASSP 2017*.

[6] P. van Dorp and F. Groen, "Human walking estimation with radar," *IET Radar, Sonar and Navigation*, vol. 150, no. 5, pp. 356–365, 2003.

[7] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Trans. Geo. and Remote Sens.*, vol. 47, no. 5, pp. 1328–1337, 2009.

[8] S. S. Ram, Y. Li, A. Lin, and H. Ling, "Doppler-based detection and tracking of humans in indoor environments," *J. Franklin Inst.*, vol. 345, no. 6, pp. 679–699, 2008.

[9] D. P. Fairchild and R. M. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-Doppler signatures," *IET Radar, Sonar and Navigation*, vol. 8, no. 5, pp. 425–434, 2014.

[10] B. Jokanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," in *IEEE Radarcon 2016*.

[11] S. Gurbuz, B. Tekeli, M. Yuksel, C. Karabacak, A. Gurbuz, and M. Guldogan, "Importance ranking of features for human micro-doppler classification with a radar network," in *FUSION*, July 2013, pp. 610–616.

[12] A. Balleri, K. Chetty, and K. Woodbridge, "Classification of personnel targets by acoustic micro-Doppler signatures," *IET Radar, Sonar and Navigation*, vol. 5, no. 9, pp. 943–951, 2011.

[13] C. Clemente, L. Pallotta, A. De Maio, J. J. Soraghan, and A. Farina, "A novel algorithm for radar classification based on Doppler characteristics exploiting orthogonal pseudo-zernike polynomials," *IEEE Trans. Aerosp. Electron. Syst*, vol. 51, no. 1, pp. 417–430, 2015.

[14] D. Tahmoush and J. Silvious, "Time-integrated range-Doppler maps for visualizing and classifying radar data," in *IEEE Radarcon 2011*.

[15] A.-K. Seifert, M. G. Amin, and A. M. Zoubir, "New analysis of radar micro-Doppler gait signatures for rehabilitation and assisted living," in *IEEE ICASSP 2017*.

[16] J. Karhunen and J. Joutsensalo, "Generalizations of principal component analysis, optimization problems, and neural networks," *Neural Netw.*, vol. 8, no. 4, pp. 549–562, 1995.

[17] C. Croux and G. Haesbroeck, "Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies," *Biometrika*, vol. 87, no. 3, pp. 603–618, 2000.

[18] T.-K. Kim, J. Kittler, and R. Cipolla, "Discriminative learning and recognition of image set classes using canonical correlations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 1005–1018, 2007.

[19] A. V. Knyazev and P. Zhu, "Principal angles between subspaces and their tangents," *Arxiv preprint*, 2012.

[20] A. A. Goshtasby, "Similarity and dissimilarity measures," in *Image registration*, 2012, pp. 7–66.

[21] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[22] A. Malhi and R. X. Gao, "PCA-based feature selection scheme for machine defect classification," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 6, pp. 1517–1525, 2004.

[23] B. Jokanovic, M. Amin, F. Ahmad, and B. Boashash, "Radar fall detection using principal component analysis," in *Proc. SPIE 2016*.