EARLY ALZHEIMER DISEASE DETECTION WITH BAG-OF-VISUAL-WORDS AND HYBRID FUSION ON STRUCTURAL BRAIN MRI

Olfa Ben Ahmed [1, 2], Jenny Benois-Pineau[1], Chokri Ben Amar[2], Michelle Allard[3], Gwenaelle Catheline[3]

Universityof Bordeaux1, Laboratoire Bordelais de Recherche en Informatique(LaBRI)[1], France
olfa.ben-ahmed@labri.fr, jenny.benois@labri.fr
University of Sfax, Research Group on intelligent machine (ReGIM)[2], Tunisia
chokri.benamar@ieee.org
University of Bordeaux 2, Aquitaine Institute for Cognitive and Integrative Neuroscience[3], France
michelle.allard@chu-bordeaux.fr, gwenaelle.catheline@ch-bordeaux.fr

ABSTRACT

In this paper, we tackle the problem of recognition of Alzheimer's disease (AD) in structural MRI images using visual similarity. AD yields visible changes in the brain structures. We aim to recognize patient category such as AD, or prodromal stage of the AD called Mild Cognitive impairment (MCI), or normal control subject (NC).We use visual local descriptors and the bag of words approach on the most involved regions in AD (Hippocampus and PosteriorCingulate Cortex) in MRI images. The Content-Based Visual information retrieval (CBVIR)approach is then applied to recognize patient category. The contribution of the paper is in the fusion of visual signatures and of classification results obtained on characteristic brain regions.

Index Terms— Alzheimer, bag of words, visual similarity, hippocampus, Posterior Cingulate Cortex, CBVIR.

1. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disease characterized by a severe deterioration in cognitive function and memory loss. It represents, actually, a major public health problem. Not only biological biomarkers are used to detect AD but also structural MRI is accepted to be used as early imaging biomarker for AD detection[1][2]. Medical information on brain magnetic resonance imaging (MRI) is used in detecting abnormalities in physiological structures. Structural (MRI) measurements can detect and follow the evolution of brain atrophy which is a marker of the disease process. Therefore, it allows for diagnosis and prediction of Alzheimer's disease [3].The current trend in early Alzheimer's disease detection is towards a local hippocampus description. In fact, a large number of studies have shown that people with AD or Mild cognitive Impairment have smaller hippocampal volumes than healthy individuals [4]. Hence the image-based diagnostics of AD relies mainly on volumetric analysis. Nevertheless, the overall volumetric analysis of the hippocampus does not describe the local change of structure, which is helpful for diagnosis.

Since recently CBVIRmethods have been applied in the field of medical image indexing [5] and disease detection. MRI image comparison techniques are being created but visual features extraction methods are not being fully exploited to describe such data. One the other hand, as noted on [5], the use of the content-based image retrieval systems may help medical staff to make the diagnosis. Thus, both facts encourage the creation of the automated system which helps clinicians to extract only relevant information from the image database. Several works used visual descriptors to extract information from human brain. We refer the reader to [6][7][8][9]. In [7], the authors are focusing on integrating different kinds of information, including textual data, image visual features extracted from scan as well as direct user (doctor) input on image similarity. Features used to describe brain images are LBP and DCT. [10]uses visual image similarity to help early diagnostics of Alzheimer. [7][8] proved the usefulness of user feedback for brain image classification. In [8]CHF and SIFT visual local descriptors are computed around the hippocampus region and several schemes of image retrieval are used. Some works evaluate the suitability of the bag of words approach [11](BOVW) for automatic classification of MRI images in the case of Alzheimer disease. In [9]the authors use the scale invariant features transform (SIFT) extracted from the whole subjects' brain to classify between brain with and without AD. It was shown in [12]that the BOF, that is a direct comparison of features instead of their quantization as

in the BOVW, approach is able to describe the visual information for discriminating healthy brain from those suffering from the AD. However [9] and [12]do not addressthe Mild Cognitive Impairment (MCI) case which has become an important issue in the study of AD. From image similarity/classification point of view, this class is very much heterogeneous. It contains the MRI with rather and certain similarity with AD or NC cases. Hence, in this paper we tackle this difficult class from CBVIR perspective. Generally, we present a framework to help early diagnosis of AD from MRI using visual descriptors and the BOVW approach. This research was initiated in [8].The contribution of present research is as flowing; first, we consider visual features extraction from two brain regions; hippocampus andPosterior Cingulate Cortex (PCC). Second, it has been shown in [8] that image signature contains too match individual brain structural information which is not relevant to characterize the disease. In order to reduce the dimensionality of the signatureobtained from fusion and to retain only relevant information we apply the PCA. Finally, in order to improve the recognition performance especially in MCI class we add an early fusion schemes for two regions: Hippocampus and PCC. This early fusion is the part of a global fusion approach using classification of Cerebrospinal fluid (CSF) [8]. This is why we further call this "hybrid fusion approach". The experiments are also extended, we use in addition to ADNI data base the real cohort of patients 3D_T1_3C3E [2]. The rest of this paper is organized as follows, in section 2, framework of features extraction, signature generation and fusion is described. In section 3, experiments and results are presented and section 4 concludes the work.

2. FRAMEWORK

Our approach consists initially in brain image normalization, which is a standard step in brain image comparison. Then, Regions-of-Interest (ROIs) are extracted from the normalized images to be described by local visual descriptors and quantified in the BOVW framework.
The *early fusion* of features is then achieved by concatenating the hippocampus and the PCC signatures. We test 2 modes of early fusion: i) simple concatenation, ii) PCA applied on concatenated signatures.
Furthermore, the preliminary classification of pixels on hippocampus area intoCerebrospinal Fluid (CSF) and brain cells is realized as in [8]. The classification result is used for similarity computation thus insuring the *late fusion*. The overall fusion approach for retrieval is therefore a *hybrid fusion*. An overview of our approach is illustrated in figure 1. In the image illustrating the CSF and brain cells classification, the hippocampus region is marked with red color and selected CSF voxels are marked with green.
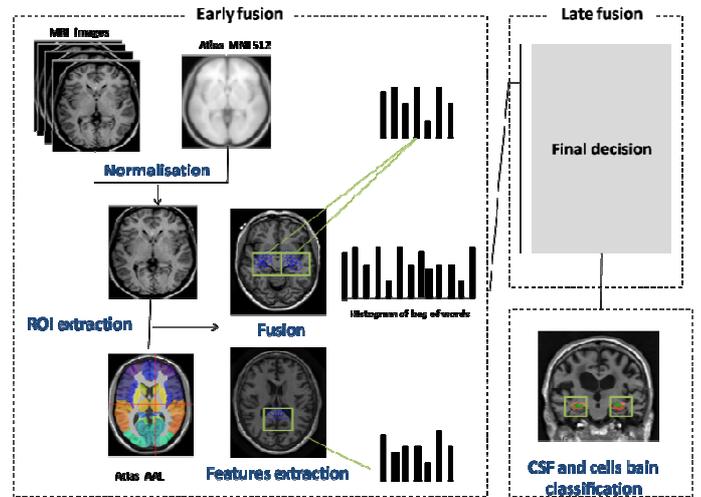

Figure 1 Framework description

2.1. Spatial normalization

Spatial normalization is the process of transforming brain images from their natural form to a standard form. It is usually done by using a standard brain called "template". Spatial normalization is an essential step for any study of groups because of large inter-individual variability of human brain. It consists in a set of transformations to reduce inter-individual variability and to facilitate subject comparison;in our work, all images are aligned to an template using linear (affine) transformations. We use the MNI 512 template which is a standard image provided by the Montreal Neurological Institute (MNI). Statistical Parametric Mapping (SPM) and the VBM toolbox are employed to fulfill the registration. We adopt only affine registration and do not use a non-linear registration to a template to retain the local structure of ROI and not to deform the features of an individual.

2.2. Features extraction

Visual information has to be extracted on the characteristics structures in the brain related to the presence of Alzheimer disease. Thus, it is necessary to locate the ROI inside the brain volume. Referring to the domain knowledge we choose to describe the regions of Hippocampus and the PCC. Indeed, in the early stage, Alzheimer's disease affects the hippocampal region that undergoes significant neuronal loss and is therefore shrinked[13][14][15]. In more advanced stages, the PCC sees its metabolism decreased and eventually shrinks[16][17]. The PCC hypo metabolism is a prediction of rapid conversion to AD [18][19]. Hence, on the one hand, the alteration of the hippocampus and the PCC may constitute preclinical markers of Alzheimer's disease. The determination of the Region of Interest (ROI) is the key for region-based analysis methods. In the AD-related

research segmentation and identification of the ROI are highly complicated and time-consuming[15][3]. In this work, we follow the approach proposed in [20] and use the AAL atlasto extract ROI. The resulting mask is generated using the SPM software with the VBM8 toolbox. In the next step visual descriptors are computed inside these regions. In this work we flow [8] and use the same local features: scale-invariant feature transform (SIFT) descriptor [21] and Circular Harmonic Functions CHF [22][23]. Figure 2 shows detected CHF keypoints on the Hippocampus and the PCC areas in MRI brain image, key points are showed in blue.
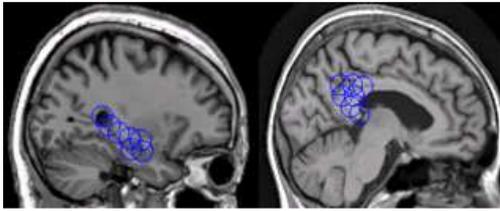

Figure 2 Key points detection on hippocampusand PCC (Sagittal projection)

We refer the reader to [22][23] for mathematical description of CHF which allow a spectral representation of image signal capturing local singularities. In [8], they proved to be more efficient than SIFT in the use of bag of features and the one to one correspondence framework.

2.3. Bag-of-Visual-Words and Hybrid fusion for similarity matching

In our work, we treat each ROI as a set of local features. Hence, theBOVW approach model is applied separately on the two ROIs (hippocampus and PCC). The role of BOVW model is to cluster extracted features from every ROI in order to build a visual vocabulary. Moreover, the region's shape differs from one projection to another. Thus, we choose to perform the clustering procedure 3 times from different projections (sagittal, axial and coronal) and to generate one visual vocabulary per projection. Firstly, all features$f_{n,i}^s$, here n and i stand respectively for slice and feature indexes, are extracted from the ROI on all slices for the sagittal projection then features are quantified by k-means algorithm. The centers $c_k^s, k \in [1, K]$ are then calculated, where K is the codebook size given as parameter to the k-means algorithm. The same is done for axial and coronal projections. All features $f_{n,i}^s, f_{n,i}^a, f_{n,i}^c$ and centers c_k^s, c_k^a, c_k^chere have the same dimensionality of the descriptor being used. In case of SIFT it is 128 and for CHF it is 150. Once the cluster centers have been determined, the image signature per projection is generated. Each feature is assigned to closest using the distance$d(f_{n,i}^s, c^s)$, here we use the Euclidian distance. Then each projection is represented by a normalized histogram of occurrence of visual words.

The image signature per ROI h is acquired by the concatenation of the histograms from all projections$h = [h^s h^s, h^s]$. Final signature is obtained by concatenating Hippocampusand PCC signatures. To reduce the resulting image signature we use the PCA (PrincipalComponents Analysis) which is a useful mathematical technique for reducing of vector dimensionality and finding an optimal combination of variables into a smaller set [24].

For a given query scan Q (x, y,z), the features are computed using the same process resulting the h^*. Image similarity is established by comparing the signatures, lower distance means more similarity. For histogram comparison, the distance metric that has been chosen is L1:

$$d_n = D(h^*, h^n) = \sum_{i=1}^{K} |h^* - h^n| \quad (1)$$

The concatenation of histograms and application of PCA constitute the *early fusion* process. Using the Bayesian classification approach from completely different perspective, namely the presence of CSF on Hippocampus ROI, we define for the query image Q (x,y,z), the probabilities to belong to three classes P_{AD}, P_{NC}, P_{MCI} (AD, NC, MCI respectively) as it is done in [8]. The combined distance is then obtained by multiplication fusion operator:

$$d_n^{class}(-\ln(p_{classeof(n)}) * d_n) \quad (2)$$

Finally, similarity to n-th image = $1/d_n^{class}$.

This is the *late fusion* part of our hybrid fusion approach. We apply this fusion as it proved to be efficient on classification into three classes in [8].

3. EXPERIMENTS AND RESUTS

To perform the test, the ground truth data on the image similarity is needed, this information however is not available. Moreover, the similarity from medical point of view would be different for different experts. Thus, in our testing procedure we only test for correct class correspondence (AD, NC, and MCI). Indeed, the images from the same class should be more similar than images from other classes and the retrieval precision can be calculated as the percent of correct image classes in the first N retrieved images. To increase the number of experiments and precision of statistics full cross-validation is performed, (we are repeating the test for each image in our test dataset taking as the database the rest of the images). For the test of our method, we used images taken from two different image databases. First, we used the public Alzheimer's disease Neuroimaging Initiative(ADNI)[1]database. Images are T1 weighted contrast, we selected 188 images for test: 40 AD, 87 MCI and 61 NC. Second, a 3T weighted contrast images

---

[1]www.loni.ucla.edu/ADNI

from the real cohort "3D_T1_3C3E" were also tested. The subset of the latter database contains 177 images (10 AD and 167 NC)" obtained from a subsample of a large French epidemiological study, the 3C-study. [2].

3.1. AD-patients versus Normal controls using only hippocampus ROI

In the first part of experiments, we evaluated the BOVW approach on the hippocampus area to distinguish between AD and NC. For ADNI data, we varied the code book size from 50 to 400. The best recognition rate was obtained for a codebook size of 210, obtained result are plotted in figure 3.
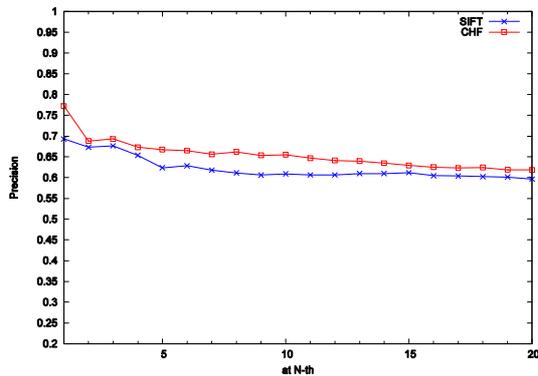

Figure 3 Precision at Nth: ADNI data

The performance, presented by precision at 1, of the BOVW reaches 77.2% in the case of only 2 classes (NC and AD) compared to [8] where performance of 74% were obtained with the CHF 1 to 1 and 68 % using BOF [8]. For the SIFT descriptor we reach 69% which is better than retrieval rate reached by SIFT 1 to 1 (64%) [8].The figure 4 shows precision for CHF and SIFT descriptors on the INCIA dataset for a codebook size of 150.
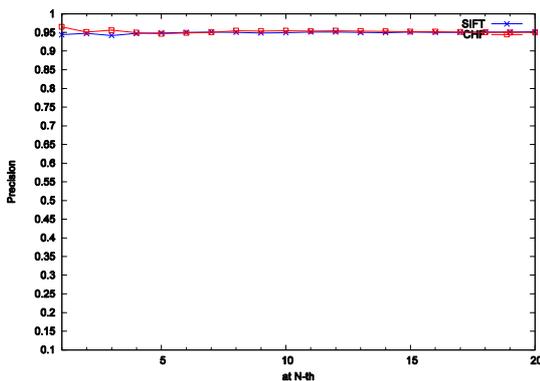

Figure 4 : Precision at Nth: 3C dataset, only hippocampus ROI

Precision reaches 96.5% for CHF descriptors and 95.17% for SIFT descriptors. In most cases, the CHF descriptor performance is slightly better. The bag of words approach performs well.

3.2. MCI retrieval using Hippocampus and PCC ROIs features fusion

The MCI category is the most difficult to recognize, as the structural changes in the characteristic brain regions are very unequal. In the case of 3 classes retrieval, performance using only hippocampus drops to 37% for SIFT descriptor, the most likely reasons are that MCI class is a transition between AD and NC thus the bounds are uncertain. As we mentioned in the previous sections, the PCC alteration can be a predictive biomarker of rapid conversion to AD and hence should characterize the MCI cases.Therefore, we extract features from both hippocampus and PCC areas. In the first BOVWexperiment, we applyan early fusion approach and concatenate hippocampus and PCC signatures in global BOVW from 3 projections and both regions are thus concatenated in the description space. The most relevant precision was obtained with a codebook K=350 yielding a signature size of 2100: (K * projections number * 2).In the second experiment we used the PCA technique to reduce the signature size to 69 for CHF and to 24 for SIFT. These dimensions were obtained experimentally by varying the number of principal components on ADNI dataset.
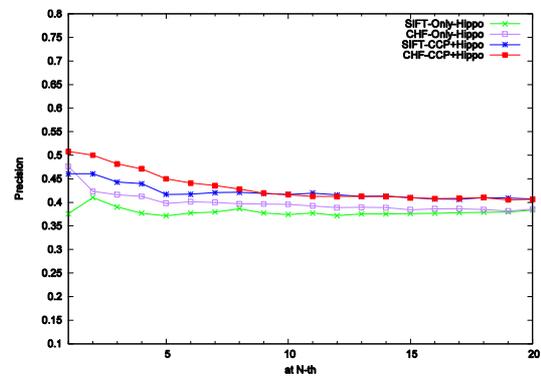

Figure 5 : Precision at Nth with and without PCA concatenation

In this case, the best precision was obtained for smaller codebook size of 150; the performance was higher by 4% for PCA on concatenated signatures compared to early signatures fusion without PCA.

As shown in figure 5, the precision of the descriptor based retrieval when using PCC and Hippocampus signatures fusion is greater than precision when using only hippocampus area. Fusion strategy increases results by an average of 6% while the fusion by dark pixel volume primary classification proposed in[8] gave only 5% of

amelioration to only hippocampus classification. From table 1 it can be seen that:

- The performance of image retrieval is substantially improved by 10% using the concatenating of PCC and hippocampus signatures and the CSF volumes classification, proposed in [8], compared to classification using only hippocampus.
- Obtained results exceed the reported results in [8] at 10th as well at 20th which are the more challenging.
- When using our early fusion approach to describe images (CHF case), retrieval precision is about 0.516 while when using only preliminary classification, precision drops to 0.45 [8].

Table 1 : Average precision at Nth using only hippocampus features, fusion of both hippocampus and PCC fusion signatures and fusion of proposed approach with the CSF volume classification, ADNI

Average precision SIFT CHF	Only hippocampus	Fusion PCC + hippocampus	Fusion PCC - Hippo + CSF volume classification
at 1 SIFT/CHF	0.3756 0.4561	0.4603 **0.5179**	**0.5080** **0.5132**
at 3 SIFT/CHF	0.3897 0.4162	0.4426 0.4810	**0.4868** **0.4840**
at 10 SIFT/CHF	0.3740 0.3957	0.4169 0.4197	**0.4571** **0.4626**
at 20 SIFT/CHF	0.3735 0.3854	0.4071 0.4111	**0.4367** **0.4421**

Compared to the results presented in [7], the precision at 10 and 20 obtained on a subset of this database are higher in the case of our full fusion schemes. In table 1, we have 0.462 (column 4 line 3) versus 0.42 and 0.442 (column 4 line 4) versus 0.38. Although we used only a pure visual image description,we obtain better results than [7] in which authors combine visual features with textual data.

4. CONCLUSION AND FUTURE WORKS

Using visual similarity between MRI images allowed us to provide the clinicians with semantic similarity, and thus could potentially support their diagnostic decision. PCC and Hippocampus features fusion improve accuracy on MCI case retrieval. Across a range of tests, useful level of recognition rates were achieved with a small signatures sizes for both CHF and SIFT descriptors. By increasing the number of subjects and using a supervised learning approach phase we can improve the performance of our approach. This is the subject of our future work.

4. REFERENCES

[1] G. Frisoni, N. Fox, C. Jack, P. Scheltens and P. Thompson, "The clinical use of structural IRM in Alzheimer disease," *Nat Rev Neurol,* vol. 6, pp. 67-77, 2010.

[2] G. Catheline, O. Periot, M. Amirault, M. Braun, J. Dartigues, S. Auriacombe and M. Allard, "Distinctive alterations of the cingulum bundle during aging and Alzheimer's disease.," *Neurobiol Aging,* 2008.

[3] C. Rémi, G. Emilie, T. Jérôme, A. Guillaume and L. Stéphane, "Automatic classification of patients with Alzheimer's disease from structural MRI, A comparison of ten methods using the ADNI database," *neuroimage,* pp. 766-780, 2011.

[4] V. Nicolas, D. Béatrice, V. Fausto, D. L. S. Vincent, M. Florence and L. Brigitte, "Relationships between hippocampal atrophy, white matter disruption, and gray matter hypometabolism in Alzheimer's disease," *Neuroscience,* 2008.

[5] P. Ghosh, S. Antani, L. Long and G. Thoma, "Review of medical image retrieval systems and future directions," in *24th International Symposium on Computer-Based Medical Systems (CBMS)*, 2011.

[6] B. Ridha, J. Barnes, J. Bartlett, A. Godbolt, T. Pepple, M. Rossor and al, "Tracking atrophy progression in familial Alzheimer's disease: a serial MRI study," *lancet neurology,* pp. 828-34, 2006.

[7] M. Agarwal and J. Mostafa, "Content-based image retrieval for Alzheimer's disease detection," in *9th International Workshop on Content-Based Multimedia Indexing (CBMI)*, June 2011.

[8] M. Mizotin, J. Benois-Pineau, M. Allard and G. Catheline, "feature based brain MRI retrieval for alheimer disease diagnosis," in *ICIP*, 2012.

[9] M. Daliri, "Automated diagnosis of Alzheimer Disease using Scale Invariant feature transform in magnectic resonance Images," *Journal of medical systems,* pp. 995-1000., 2011.

[10] B. A. Ceyhun and N. Eindhoven, "Automated diagnosis of Alzheimer's disease using image similarity and user feedback," in *International Conference on Image and Video Retrieval*, 2009.

[11] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," in *In Proc of 9 th IEEE Conf on Computer Vision* , 2003.

[12] A. Rueda, J. Arevalo, A. Cruz, E. Romero and F. González, "Bag of Features for Automatic Classification of Alzheimer's Disease in Magnetic Resonance Images," *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications,* pp. 559-566, 2012.

[13] B. Gutman, Y. Wang, J. Morra, A. Toga and P. Thompson, "Disease Classification with Hippocampal Shape Invariants," *Hippocampus,* vol. 19, p. 572–578, 2009.

[14] B. H. Ridha, J. Barnes, L. A van de Pol, J. M Schott, R. G Boyes, M. M Siddique, M. N Rossor, P. Scheltens and N. C Fox, "Application of automated medial temporal lobe atrophy scale to Alzheimer disease.," *Arch Neurol,* pp. :849-854, 2007.

[15] G. E, G. Chételat, M. Chupin, R. Cuingnet, B. Desgranges, H. Kim, M. Niethammer, B. Dubois, S. Lehéricy, L. Garnero, E. Eustache, O. Colliot and t. A. D. N. The Initiative, "Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging," *Neuroimage,* 2009.

[16] B. S. G. F. C. C. S. J. Callen DJ, "Beyond the hippocampus: MRI volumetry confirms widespread limbic atrophy in AD," *Neurology,* p. 57:1669, 2001.

[17] D. B. L. B. M. F. P. J. d. l. S. V. Chételat G, "Direct voxel-based comparison between grey matter hypometabolism and atrophy in Alzheimer's disease," *Brain,* p. 131:60–71, 2008.

[18] G. Chetelat, B. Desgranges, V. de la Sayette, F. Viader, F. Eustache and J. Baron, " Mild cognitive impairment: Can FDG-PET predict who is to rapidly convert to Alzheimer's," *Neurology ,* vol. 60, pp. 1374-1377, 2003.

[19] P. Nestor, T. Fryer, M. Ikeda and J. Hodges, "Retrosplenial cortex (BA 29/30) hypometabolism in mild cognitive impairment (prodromal Alzheimer's disease).," *Eur J Neurosci ,* vol. 18, pp. 2663- 2667, 2003.

[20] A. John and J. Karl, "Unified segmentation," *Neuroimage,* vol. 26, no. 3, pp. 839-851, 2005.

[21] D. Lowe, "Distinctive Image Features from Scale-Invariant keypoints," *International Journal of Computer Vision,* vol. 60, no. 2, pp. 91-110, 2010.

[22] L. Sorgi, N. Cimminiello and A. Neri, "Keypoints Selection in the Gauss Laguerre Transformed Domain," *in Proc. BMVC,* pp. 133-142, Sep. 2006.

[23] D. Sorokin, M. Mizotin and A. Krylov, "Gauss-Laguerre Keypoints Extraction Using Fast Hermite Projection Method," *Sorokin D.V., Mizotin M.M., Krylov A.S. "Gauss- Laguerre Keypoints Extra Lecture Notes in Computer Science,* vol. 6753, pp. 284-293, 2011.

[24] I. Jolliffe, "Principal Component Analysis," *2nd Edn, New York: Springer-Verlag ,* 2002.