Detection of Salient Image Points using Principal Subspace Manifold Structure

António R. C. Paiva[1] and Tolga Tasdizen[1,2]
[1] *Scientific Computing and Imaging Institute,* [2] *Dept. of Electrical and Computer Eng.,*
University of Utah, Salt Lake City, UT 84112
email: {*arpaiva,tolga*}*@sci.utah.edu*

***Abstract*—This paper presents a method to find salient image points in images with regular patterns based on deviations from the overall manifold structure. The two main contributions are that: (i) the features to extract salient point are derived directly and in an unsupervised manner from image neighborhoods, and (ii) the manifold structure is utilized, thus avoiding the assumption that data lies in clusters and the need to do density estimation. We illustrate the concept for the detection of fingerprint minutiae, fabric defects, and interesting regions of seismic data.**

***Keywords*-salient image points; manifold learning; manifold of image neighborhoods.**

I. Introduction

Detection of salient image points is an important processing step in many image analysis and computer vision applications [1], [2], [3]. Generally speaking, salient image points correspond to regions in an image that standout with regards to its context, thus calling for further attention. Note that, unlike other interest point detectors [4], in this work it does not suffice that the image signal changes two-dimensionally since that perspective would not work for general textures. This distinction can be observed, for example, in fabric images where salient points may correspond to fabric defects [5], but the image signal changes two-dimensionally almost everywhere.

In many applications, salient image point detectors are important because, once the salient image points or regions are found, the problem can be reduced to that of finding a correspondence between these points. Fingerprint matching, for example, can be achieved by finding a correspondence between salient points called minutiae [6]. A directly related problem is that of finding salient features; that is, features that indicate salient points [3]. The method presented here addresses both problems.

An important salient point detector approach, especially in texture inspection, requires the design of descriptors that characterize the image structure [7], [1], [8]. Then, the salient points are detected by computing the descriptors in sliding windows and compare their values to those in a reference window. The primal disadvantages of this approach are that the characterization ability is constrained by the design of the descriptors and the need to establish a reference feature pattern.

Another approach is based on the idea that salient points contain the most relevant information. This means that salient points are prominent in spite of their low probability. For this reason, several image salient point detection and feature extraction methods utilize outlier detection techniques [2]. The fundamental difficulty of this perspective is that it requires estimation of the underlying probability distribution. Although several numerical estimation algorithms have been proposed, these methods assume that the data points form clusters [2], which may not hold in many cases, as it is shown in this paper. Moreover, the salient features are typically selected from a pre-specified family of features to avoid handling high-dimensional spaces, where the estimation of the distribution is more challenging.

In contrast, this paper presents a method for finding salient image points in textured images by finding deviations from the regular manifold of image neighborhoods. By using the manifold structure of image neighborhoods, the underlying low-dimensionality of the manifold can be utilized to avoid estimating the distribution and avoids the assumption of data clusters. Consequently, salient features are found directly and in an unsupervised manner from the space of image neighborhoods by using dimensionally reduction.

II. Manifold of image neighborhoods

A manifold is a subspace that is locally Euclidean of dimension n, with n typically much smaller than the dimension of the ambient space where the manifold is embedded. To obtain this representation, image points can be embedded in a high-dimensional space by forming a vector from image neighborhoods, or "patches," centered at the pixel location. For an image I, an $M \times M$ patch corresponding to image point (i, j) is the ordered set of points $\mathbf{x} = \{I(u, v) : |u - i| \leq (M-1)/2 \wedge |v - j| \leq (M-1)/2\}$. This representation has the advantage that is consistent with Markov random field (MRF) theory since the context information needed to completely characterize the joint distribution is preserved [9].

Consider an image with a regular pattern, and its points embedded using image neighborhoods. In the feature space, points with similar neighborhoods will be close to each other, and farther from dissimilar ones. Moreover, due to correlations commonly observed between nearby image neighborhoods, one can expect that the transitions between groups of very different patches to be somewhat smooth, and that these transitions are controlled by some underlying characteristic

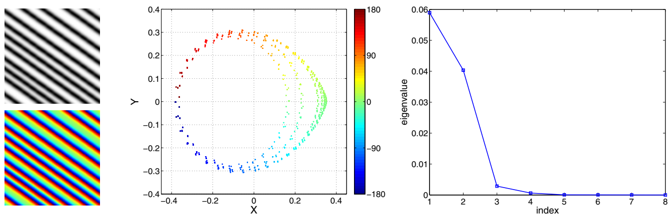

Fig. 1: Projection of the manifold of 5×5 image neighborhoods. Left: Image from which the patches where obtained (top), and phase image of the projected points (bottom). Middle: Patches projected onto 2-D color coded by phase. The colors match those in the phase image. Right: First 8 PCA eigenvalues versus the number of embedding dimensions.

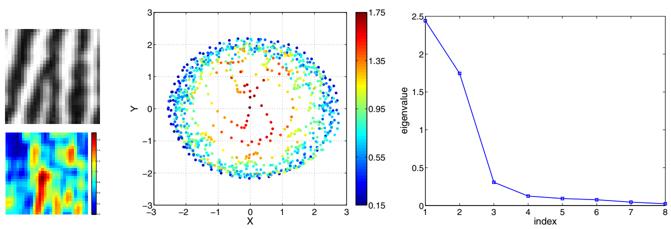

Fig. 2: Projection of the manifold of 7×7 image neighborhoods of an image containing a salient point. Sub-figures are as in Fig. 1, but color coded with regards to the norm of the projection error vector.

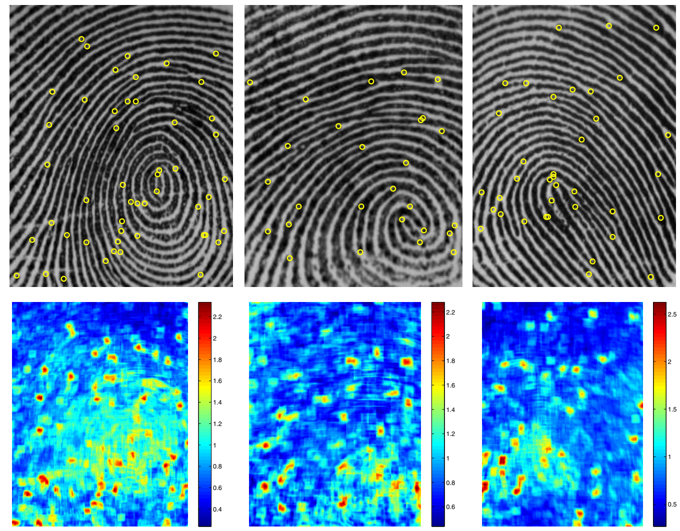

Fig. 3: Detection of minutiae in fingerprints. The original fingerprint images with the minutiae marked are shown the first row and the corresponding projection error images are shown in the second row.

or feature. These are the same principles at the core of manifold learning and nonlinear dimensionality reduction methods. These ideas are illustrated in Fig. 1, using PCA for dimensionally reduction. Note the circular shape of the manifold, which clearly violates the assumption of data clusters made by some methods.

The importance of these ideas for salience detection arises from the fact that salient image points do not lie in the manifold of image neighborhoods in regular regions, i.e., they are outliers. As reviewed earlier, previous approaches are based on an estimate of the distribution. However, we avoid this process by utilizing the intrinsic low-dimensionally of the manifold of regular image neighborhoods. This means that salient points will not be well characterized by the projection subspace and have large projection error (cf. Fig. 2). Note that this observation is true even if data points do form clusters. Using PCA, the projection error for patch $\mathbf{x}_i$ is defined as

$$p_e(\mathbf{x}_i) = \left\| \mathbf{x}_i - \mathbf{V}\mathbf{V}^T \mathbf{x}_i \right\|, \qquad (1)$$

where $\mathbf{V}$ is the matrix with the PCA projection column eigenvectors.

A fundamental advantage of the methodology we propose is that the salient features are determined directly in an unsupervised manner and depend only on the data due to the use of dimensionality reduction methods. This is because dimensionality reduction algorithms must find the data subspace and an appropriate projection that preserves the underlying manifold structure. For instance, this allows in the example in Fig. 1 for a projection into two dimensions[1] to be utilized, due to having only two high eigenvalues, even though the *image pattern is aperiodic*, which would pose a problem to methods with fixed features or descriptors.

Using the manifold structure for salience detection involves setting three parameters: the patch size, the threshold, and the number of principal components in the case of PCA, or equivalent parameter in the case of another dimensionality reduction method. The choice of patch size is not critical as long as it is comparable to the period of the texture. In practice, however, it is helpful to carefully choose the smallest possible image neighborhood size that achieves reliable results in order to reduce the computational complexity, which is $\mathcal{O}(Nd^2 + d^3)$, where N is the number of data samples and d the dimensionality of the feature vector. Typically, the threshold can be easily estimated from a few example images. The number of principal components to use will depend on the application. In most cases, it is fixed and known a priori, as in Section III-A, or can be determined automatically from data, as shown in Section III-B.

III. Example applications

A. Fingerprint minutiae detection

The first application example in on the detection of fingerprint minutiae. Fingerprint minutiae are points where the ridge pattern at the fingertip is discontinuous. The two most common cases of minutiae are points where the ridge terminates or bifurcates. Other cases include crossovers, spurs, islands and lakes; see Maltoni et al. [6] for a complete reference. Minutiae are very important because their spatial pattern is highly

[1] In this case, the manifold is a circle which is topologically 1-D but only embeddable in 2-D.

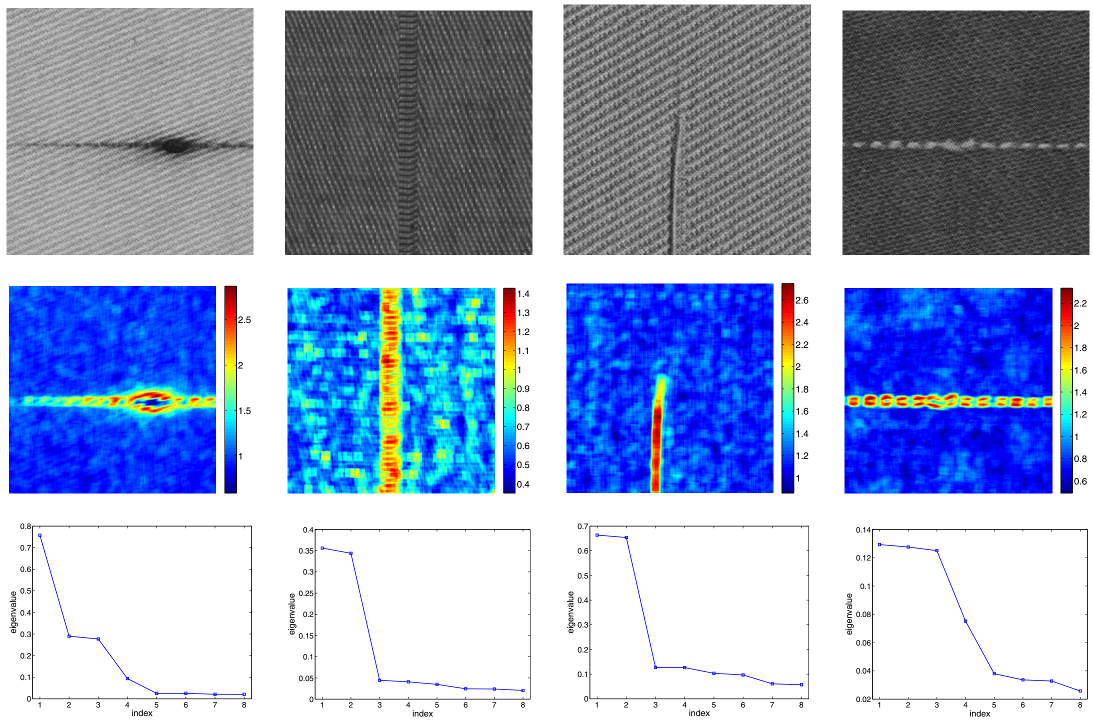

Fig. 4: Detection of fabric defects. Top: Original images. Middle: Image with the corresponding projection error. The number of principal components was determined by the number of major eigenvalues. From top to bottom, we used 1, 2, 2, and 3 principal components. Bottom: Eigenvalues for the first 8 principal components.

characteristic of each individual and, consequently, most fingerprint identification systems match fingerprints simply by finding a correspondence between minutiae points.

Unlike the case used in the previous section, however, the overall manifold of fingerprint patches has a much more convolved structure due to the different orientations of the ridge pattern. For that reason, fingerprints are said to be *quasi*-regular. Consequently, one approach would be to utilize nonlinear dimensionality reduction methods, but for image analysis the computational complexity of these methods is extremely high, which is $\mathcal{O}(N^2)$ with N the number of patches. It is easy to see that even for relatively small images N will easily be on the order of several tens of thousands, and therefore much larger than d^2. Hence, instead, we approximate the manifold locally by applying PCA analysis in overlapping windows. Basically, because the ridge orientation is approximately constant within the window, the points lie locally in a linear 2-D subspace in feature space. Then, the projection error can be obtained using Eq. 1.

The results on three fingerprints of the Fingerprint Verification Competition (FVC 2002) [6] are shown in Fig. 3. For simplicity and to avoid the problem of fingerprint segmentation, the fingerprint images were cropped to ensure that the whole image contained only the fingerprint region. The PCA analysis was done in 18×18 windows overlapping by 12 pixels in either direction, and using 11×11 image neighborhoods. The projection error was computed with regards to the projection onto the first two principal components. To find the minutiae points, the projection error image was thresholded at 1.7 in this case, and for each connected component a minutia point was marked at the location with the largest error. The results in Fig. 3 show that the method described here successfully marks nearly all true minutiae, albeit it still detects a few spurious minutiae, most likely due to our simplistic thresholding approach.

B. *Fabric defect detection*

The second example is on the detection of fabric defects [10]. The ability to detect fabric defects is key for quality control in the textile industry. In this process, an automated inspection system is very helpful because it can reduce production costs by allowing higher yields due to the high analysis speeds [5].

We applied our approach to some of the image utilized in Mak et al. [5]. Unlike the previous example, because the fabric textures are overall regular, in this case we applied PCA globally for dimensionality reduction. Note that we are not assuming that points form a cluster, only that defects are outliers in an orthogonal direction.

The results on different fabrics and with different type of defects are shown in Fig. 4. We used 15×15 image neighborhoods, corresponding to 225-dimensional feature vectors. Note the difference in the distribution of the eigenvalues for different fabric textures, due to the differences in the pattern. This difference was utilized to *automatically* determine the number of principal components to utilize. Only the principal components before the largest decrease in eigenvalue were considered. This is demonstrated in the middle column of

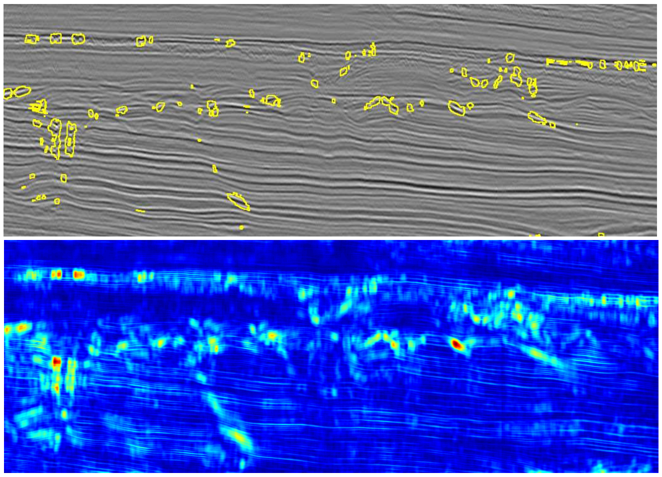

Fig. 5: Detection of interesting regions in seismic data. Top: Original image with the detected regions marked. Bottom: The corresponding projection error image.

Fig. 4, where the projection error image shown was computed according to this approach. It is important to emphasize that the method handles differences in pattern, contrast and mean intensity naturally through the adaptation of its projection basis, as mentioned earlier. The only parameter that must be specified by the user is the size of the image neighborhoods, which defines the feature vector. The value used in this particular example was chosen conservatively to ensure reliable results for all images shown. To apply this approach for defect detection, one needs also to select the threshold on the projection error above which a point is marked as a defect with regards to a desired sensitivity. However, this parameter can easily be estimated from a few example images.

C. Seismic data analysis

Our last example is on the detection of potentially interesting geological regions in seismic data. In several applications, exploration seismology is utilized to find archaeological sites, mineral deposits, and to obtain geological information for civil construction [11]. In these studies, geologists collects large volumes of data that are utilized to characterize the terrain. These volumes are then painstakingly analyzed by an expert to identify regions of interest. Instead, we propose to utilize the methodology described in Section II to analyze the volume and suggest potentially interesting regions where an expert can focus their attention.

For this experiment we utilized the 'Fortescue' seismic volume available in the public-domain from the Australian government. Again, we applied PCA globally due to the overall regularity of the layered structure. The results using PCA and 11×11 image neighborhoods on a section of the volume are shown in Fig. 5. The projection error is measured with regards to the projection onto the first four principal components, chosen as in the previous example. Although we show results on only one section, the same methodology has been applied to the whole volume using 3-D neighborhoods with similar results. From the results in Fig. 5 it can be verified that our approach marks regions where the regular layered pattern of the seismic data is disrupted. The threshold on the projection error image was set empirically.

IV. Conclusion

This paper demonstrates the use of manifold structure for salience detection in images. The fundamental observation is that the manifold of image neighborhoods from an image with a regular pattern lies in a low dimensional subspace and salient points break this pattern. Therefore, salient image points fall outside of this subspace, and have large reconstruction error. The key advantage compared to application specific methods is that the data subspace is determined directly and without supervision by the dimensionality reduction algorithm. Hence, the method proposed is general and applicable in multiple situations without methodological differences, as exemplified in three different applications. It must be remarked that the application examples were meant only to illustrate the generality of the concept proposed, rather than proposing new approaches to those problems.

One of the characteristics in utilizing the manifold structure to infer information about the image is that it shares similarities to visual systems. This is because visual systems are capable of extracting salient features just by contrasting these points with their local context and/or with the perception of the image structure at a broader level [12]. In a sense, this is very similar to what we obtain by embedding the image points and computing locally the manifold topology.

References

[1] N. Sebe, Q. Tian, E. Louplas, M. S. Lew, and T. S. Huang, "Evaluation of salient point techniques," *Image and Vision Computing*, vol. 21, no. 13–14, pp. 1087–1095, Dec. 2003.

[2] D. Lisin, E. Riseman, and A. Hanson, "Extracting salient image features for reliable matching using outlier detection techniques," in *Proceedings of the International Conference on Computer Vision Systems*, Graz, Austria, Apr. 2003, pp. 481–491.

[3] J. H. Piater, "Visual feature learning," Ph.D. dissertation, University of Massachusetts Amherst, Feb. 2001.

[4] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 151–172, Jun. 2000.

[5] K. Mak, P. Peng, and K. Yiu, "Fabric defect detection using morphological filters," *Image and Vision Computing*, vol. 27, no. 10, pp. 1585–1592, Sep. 2009.

[6] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*. New York, NY: Springer, 2003.

[7] T. S. Newman and A. K. Jain, "A survey of automated visual inspection," *Computer Vision and Image Understanding*, vol. 61, no. 2, pp. 231–262, Mar. 1995.

[8] D. Chetverikov and K. Gede, "Textures and structural defects," in *Proceedings of the International Conference on Computer Analysis of Images and Patterns*, 1997, pp. 167–174.

[9] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. Springer, 2001.

[10] R. Meylani, A. Ertüzün, and A. Erçil, "2-D robust recursive least squares lattice algorithm for defect inspection of textile products," in *Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, vol. 2, Antalya, Turkey, Jun. 1999, pp. 593–597.

[11] R. Sheriff and L. Geldart, *Exploration Seismology*, 2nd ed. Cambridge University Press, 1995.

[12] R. Linsker, "Self-organization in a perceptual network," *Computer*, vol. 21, no. 3, pp. 105–117, 1988.