E. A. Engel

GRAPHIC INFORMATION PROCESSING USING INTELLIGENT ALGORITHMS

Finding an appropriate set of features is an essential problem in the design of a shape recognition system. This paper attempts to show that for recognition of objects with high shape variability such as handwritten characters and human faces it is preferable to use the modified artificial neural network to feed the system with processed images by novel scale- and rotation-invariant interest point detectors and descriptors and to rely on learning to extract the right set of features. Experiments have confirmed the usefulness of the modified artificial neural network in a real-world application.

Keywords: object recognition, detector-description-modified artificial neural network scheme.

The task of finding correspondences between two images of the same scene or object is part of many computer vision applications. Camera calibration, 3D reconstruction, image registration, and object recognition are just some of them. The search for discrete image correspondences can be divided into three main steps. First, "interest points" are selected at distinctive locations in the image, such as corners, blobs, and T-junctions. The most valuable property of an interest point detector is its repeatability, i. e. whether it reliably finds the same interest points under different viewing conditions. Next, the neighborhood of every interest point is represented by a feature vector. This descriptor has to be distinctive and, at the same time, robust to noise, detection errors, and geometric and photometric deformations. Finally, the descriptor vectors are matched between different images. The matching is often based on a distance between the vectors, e.g. the Mahalanobis or Euclidean distance. The dimension of the descriptor has a direct impact on the time this takes, and a lower number of dimensions are therefore desirable.

It has been our goal to choose both a detector and descriptor and develop a matching step base on modified artificial neural network, which in comparison to the state-of-the-art are faster in computing, while not sacrificing performance. In order to succeed, one has to strike a balance between the above requirements, like reducing the descriptor's dimension and complexity, while keeping it sufficiently distinctive.

A wide variety of detectors and descriptors have already been proposed in researches (e.g. [1–6]). Also, detailed comparisons and evaluations on benchmarking datasets have been performed [7–9]. While constructing our fast detector and descriptor, we built on the insights gained from this previous work in order to get a feel for what are the aspects contributing to performance. In our experiments on benchmark image sets as well as on a real object recognition application, the resulting detector and descriptor are not only faster, but also more distinctive and equally repeatable.

When working with local features, a first issue that needs to be settled is the required level of invariance. This clearly, depends on the expected geometric and photometric deformations, which in turn are determined by the possible changes in viewing conditions. Here, we focus on scale and image rotation invariant detectors and descriptors. These seem to offer a good compromise between feature complexity and robustness to commonly occurring deformations. Skew, anisotropic scaling and perspective effects are assumed to be second-order effects that are covered to some degree by the overall robustness of the descriptor. As also claimed by Lowe [2], the additional complexity of full affine-invariant features often has a negative impact on their robustness and does not pay off, unless really large viewpoint changes are to be expected. In quite a few applications, like mobile robot navigation or visual tourist guiding, the camera often only rotates about the vertical axis. The benefit of avoiding the overkill of rotation invariance in such cases is not only increased speed, but also increased discriminative power.

Related works. Our results are based on the following works.

Interest Point Detectors. The most widely used detector is, probably the Harris corner detector [10], proposed back in 1988, based on the eigenvalues of the second-moment matrix. However, Harris corners are not scale-invariant. Lindeberg introduced the concept of automatic scale selection [1]. This allowed the detection of interest points in an image, each with their own characteristic scale. He experimented with both the determinant of the Hessian matrix as well as the Laplacian (which corresponds to the trace of the Hessian matrix) to detect blob-like structures. Mikolajczyk and Schmid refined this method, creating robust and scale-invariant feature detectors with high repeatability, which they coined Harris-Laplace and Hessian-Laplace [11]. They used a (scale-adapted) Harris measure or the determinant of the Hessian matrix to select the location, and the Laplacian to select the scale. Focusing on speed, Lowe [12] approximated the Laplacian of Gaussian (LoG) by a Difference of Gaussians (DoG) filter.

Several other scale-invariant interest point detectors had been proposed. Examples are the salient region detector proposed by Kadir and Brady [13], which maximizes the entropy within the region, and the edge-based region detector proposed by Jurie et al. [14]. They seem less amenable to acceleration though. Also, several affine-invariant feature detectors have been proposed that can cope with longer viewpoint changes. However, these fall outside the scope of this paper.

By studying the existing detectors and from published comparisons [15; 8], we can conclude that (1) Hessian-based detectors are more stable and repeatable than their Harris-based counterparts. Using the determinant of the Hessian matrix rather than its trace (the Laplacian) seems more advantageous, as it fires lesson elongated, ill-localized structures. Also, (2) approximations like the DoG can bring speed at a low cost in terms of lost accuracy.

Feature Descriptors. An even larger variety of feature descriptors has been proposed, like Gaussian derivatives [16], moment invariants [17], complex features [18; 19], steerable filters [20], phase-based local features [21], and descriptors representing the distribution of smaller-scale features within the interest point neighborhood. The latter, introduced by Lowe [2], have been shown to outperform the others [7]. This can be explained by the fact that they capture a substantial amount of information about the spatial intensity patterns, while at the same time being robust to small deformations or localization errors. The descriptor in [2], called SIFT for short, computes a histogram of local oriented gradients around the interest point and stores the bins in a 128-dimensional vector (8 orientation bins for each of the 4×4 location bins).

Various refinements on this basic scheme have been proposed. Ke and Sukthankar [4] applied PCA on the gradient image. This PCA-SIFT yields a 36-dimensional descriptor which is fast for matching, but proved to be less distinctive than SIFT in a second comparative study by Mikolajczyk et al. [8] and slower feature computation reduces the effect of fast matching. In the same paper [8], the authors have proposed a variant of SIFT, called GLOH, which proved to be even more distinctive with the same number of dimensions. However, GLOH is computationally more expensive.

The SIFT descriptor still seems to be the most appealing descriptor for practical uses, and hence also the most widely used nowadays. It is distinctive and relatively fast, which is crucial for on-line applications. Recently, Se et al. [22] implemented SIFT on a Field Programmable Gate Array (FPGA) and improved its speed by an order of magnitude. However, the high dimensionality of the descriptor is a drawback of SIFT at the matching step. For on-line applications on a regular PC, each one of the three steps (detection, description, matching) should be faster still. Lowe proposed a best-bin-first alternative [2] in order to speed up the matching step, but this leads to lower accuracy.

Approach. We use a novel detector-descriptor scheme and modified artificial neural network not only for matching step but also for classification and recognition. The detector is based on the Hessian matrix [11; 1], but uses a basic approximation; just as DoG [2] is a basic Laplacian-based detector. It relies on integral images to reduce the computation time and we therefore call it the "Fast-Hessian" detector. The descriptor, on the other hand, describes a distribution of Haar-wavelet responses within the interest point neighborhood. Again, we exploit integral images for speed. Moreover, only 64 dimensions are used, reducing the time for feature computation and matching, and increasing simultaneously the robustness. The matching is carried out using a modified artificial neural network, which increases not only the matching speed, but also the robustness of the descriptor.

In order to make the paper self-contained, we succinctly discussed the concept of integral images, as defined by [23]. They allow the fast implementation of box type convolution filters. The entry of an integral image $I_\Sigma(X)$ at a location $X = (x, y)$ represents the sum of all pixels in the input image I of a rectangular region formed by the point X and the origin,

$$I_S(X) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i, j).$$ With I_Σ calculated, it only takes four

additions to calculate the sum of the intensities over any upright, rectangular area, independent of its size.

The Fast-Hessian Detector. Our detector is based on the Hessian matrix because of its good performance in computing time and accuracy. However, rather than using a different measures for selecting the location and the scale (as was done in the Hessian-Laplace detector [11]), we rely on the determinant of the Hessian for both. Given a point $X = (x, y)$ in an image I, the Hessian matrix $H(x, \sigma)$ in X at scale σ is defined as follows:

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{yx}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix}, \quad (1)$$

where $L_{xx}(x, \sigma)$ is the convolution of the Gaussian second order derivative $\dfrac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point X, and similarly for $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$.

Gaussians are optimal for scale-space analysis, as shown in [24]. In practice, however, they need to be discredited and cropped (Fig. 1 left half), and even with Gaussian filters aliasing still occurs as soon as the resulting images are sub-sampled. Also, the property that no new structures can appear while going to lower resolutions may have been proven in the 1D case, but are known to not apply in the relevant 2D case [25]. Hence, the importance of the Gaussian seems to have been somewhat overrated in this regard, and here we test a simpler alternative. As Gaussian filters are not ideal in any case, and given Lowe's success with LoG approximations, we push the approximation even further with box filters (Fig. 1 right half). These approximate second order Gaussian derivatives, and can be evaluated very fast using integral images, independently of size.
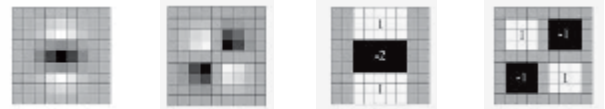

Fig. 1. Left to right: the (discredited and cropped) Gaussian second order partial derivatives in *y*-direction and *xy*-direction, and our approximations thereof using box filters. The grey regions are equal to zero

The 9×9 box filters in Figure 1 are approximations for Gaussian second order derivatives with $\sigma = 1.2$ and represent our lowest scale (i. e. highest spatial resolution). We denote our approximations by D_{xx}, D_{yy}, and D_{xy}. The weights applied to the rectangular regions are kept simple for computational efficiency, but we need to further balance the relative weights in the expression for the

Hessian's determinant with $\dfrac{\left|L_{xy}(1,2)\right|_F}{\left|L_{xx}(1,2)\right|_F} \dfrac{\left|D_{xx}(9)\right|_F}{\left|D_{xy}(9)\right|_F}$, where $|x|_F$ is the Frobenius norm. This yields

$$\det(H_{\text{approx}}) = D_{xx}D_{yy} - (0{,}9D_{xy})^2. \quad (2)$$

Furthermore, the filter responses are normalized with respect to the mask size. This guarantees a constant Frobenius norm for any filter size.

Scale spaces are usually implemented as image pyramids. The images are repeatedly smoothed with a Gaussian and subsequently sub-sampled in order to achieve a higher level of the pyramid. Due to the use of box filters and integral

images, we do not have to iteratively apply the same filter to the output of a previously filtered layer, but instead can apply such filters of any size at exactly the same speed directly on the original image, and even in parallel (although the latter is not exploited here). Therefore, the scale space is analyzed by up-scaling the filter size rather than iteratively reducing the image size. The output of the above 9×9 filter is considered as the initial scale layer, to which we will refer as scale $s = 1.2$ (corresponding to Gaussian derivatives with $\sigma = 1.2$). The following layers are obtained by filtering the image with gradually bigger masks, taking into account the discrete nature of integral images and the specific structure of our filters. Specifically, this results in filters of size 9×9, 15×15, 21×21, 27×27, etc. At larger scales, the step between consecutive filter sizes should also scale accordingly. Hence, for each new octave, the filter size increase is doubled (going from 6 to 12 to 24). Simultaneously, the sampling intervals for the extraction of the interest points can be doubled as well.

As the ratios of our filter layout remain constant after scaling, the approximated Gaussian derivatives scale accordingly. Thus, for example, our 27×27 filter corresponds to $\sigma = 3 \times 1.2 = 3.6 = s$. Furthermore, as the Frobenius norm remains constant for our filters, they are already scale normalized [26]. In order to localize interest points in the image and over scales, nonmaximum suppression in a $3 \times 3 \times 3$ neighborhood is applied. The maxima of the determinant of the Hessian matrix are then interpolated in scale and image space with the method proposed by Brown et al. [27]. Scale space interpolation is especially important in our case, as the difference in scale between the first layers of every octave is relatively large.

The Descriptor. The good performance of SIFT compared to other descriptors [8] is remarkable. Its mixing of crudely localized information and the distribution of gradient related features seems to yield good distinctive power while fending off the effects of localization errors in terms of scale or space. Using relative strengths and orientations of gradients reduces the effect of photometric changes.

The proposed descriptor is based on similar properties, with a complexity stripped down even further. The first step consists of fixing a reproducible orientation a square region aligned to the selected orientation, and extract the descriptor from it. These two steps are now explained in turn.

Orientation Assignment. In order to be invariant to rotation, we identify a reproducible orientation for the interest points. For that purpose, we first calculate the Haar Wavelet responses in x and y direction, shown in Figure 2, and this in a circular neighborhood of radius $6s$ around the interest point, with s the scale at which the interest point was detected. Also the sampling step is scale dependent and chosen to be s. In keeping with the rest, also the wavelet responses are computed at that current scale s. Accordingly, at high scales the size of the wavelets is big. Therefore, we use again integral images for fast filtering. Only six operations are needed to compute the response in x or y direction at any scale. The side length of the wavelets is $4s$.

Once the wavelet responses are calculated and weighted with a Gaussian ($\sigma = 2.5s$) centered at the interest point, the responses are represented as vectors in a space with the horizontal response strength along the abscissa and the vertical response strength along the ordinate. The dominant

orientation is estimated by calculating the sum of all responses within a sliding orientation window covering an angle of $\pi/3$. The horizontal and vertical responses within the window are summed. The two summed responses then yield a new vector. The longest such vector lends its orientation to the interest point. The size of the sliding window is a parameter, which has been chosen experimentally. Small sizes fire on single dominating wavelet responses, large sizes yield maxima in vector length that are not outspoken. Both result in an unstable orientation of the interest region.
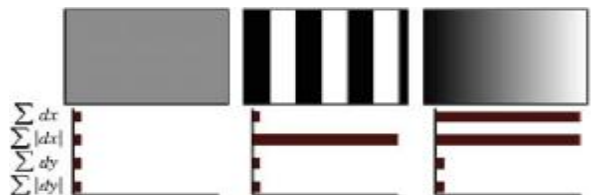

Fig. 2. The descriptor entries of a sub-region represent the nature of the underlying intensity pattern. Left: In case of a homogeneous region, all values are relatively low. Middle: In presence of frequencies in x direction, the value of $\Sigma|d_x|$ is high, but all others remain low. If the intensity is gradually increasing in x direction, both values Σd_x and $\Sigma|d_x|$ are high

Descriptor Components. For the extraction of the descriptor, the first step consists of constructing a square region centered on the interest point, and oriented along the orientation selected in the previous section. The size of this window is $20s$.

The region is split up regularly into smaller 4×4 square sub-regions. This keeps important spatial information in. For each sub-region, we compute a few simple features at 5×5 regularly spaced sample points. For reasons of simplicity, we call d_x the Haar wavelet response in horizontal direction and d_y the Haar wavelet response in vertical direction (filter size $2s$). "Horizontal" and "vertical" here is defined in relation to the selected interest point orientation. To increase the robustness towards geometric deformations and localization errors, the responses d_x and d_y are first weighted with a Gaussian ($\sigma = 3.3s$) centered at the interest point.

Then, the wavelet responses d_x and d_y are summed up over each sub-region and form a first set of entries to the feature vector. In order to bring in information about the polarity of the intensity changes, we also extract the sum of the absolute values of the responses, $|d_x|$ and $|d_y|$. Hence, each sub-region has a four-dimensional descriptor vector v for its underlying intensity structure $v = (d_x, d_y, |d_x|, |d_y|)$. This results in a descriptor vector for all 4×4 sub-regions of length 64. The wavelet responses are invariant to a bias in illumination (offset). Invariance to contrast (a scale factor) is achieved by turning the descriptor into a unit vector.

Figure 2 shows the properties of the descriptor for three distinctively different image intensity patterns within a sub-region. One can imagine combinations of such local intensity patterns, resulting in a distinctive descriptor.

In order to arrive at these descriptors, we experimented with fewer and more wavelet features, using d_x^2 and d_y^2, higher-order wavelets, PCA, median values, average values, etc. From a thorough evaluation, the proposed sets turned

out to perform best. We then varied the number of sample points and sub-regions. The 4 × 4 sub-region division solution provided the best results. Considering finer subdivisions appeared to be less robust and would increase matching times too much.

Experimental Results. The modified neural network [28] solves practical tasks in various subject fields. To investigate the generality of the detector-descriptor-modified artificial neural network scheme we solved object recognition tasks (handwritten digits, human faces). This section reports the results of numerical experiments which indicate, that the modified neural network with detector-descriptor scheme as preprocessing step has appropriate generalization accuracy.

MNIST. The MNIST database of handwritten digits has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger set available from NIST. The digits have been size-normalized and centered in a fixed-size image. The 60,000 pattern training set contains examples from approximately 250 writers.

Many methods had been tested with this training set and test set (Table 1). Some of those experiments used a version of the database where the input images where deskewed (by computing the principal axis of the shape that is closest to the vertical, and shifting the lines so as to make it vertical). In some other experiments, the training set was augmented with artificially distorted versions of the original training samples. The distortions are random combinations of shifts, scaling, skewing, and compression. First we obtain database using detector-descriptor scheme on the MNIST database. For detector and descriptor we using the testing software provided by Mikolajczyk (URL : http://www.robots.ox.ac.uk/vgg/research/affine/). The matching is carried out as follows. There were 1 000 2-layer modified artificial neural networks (33 hidden units: 30 and 3 units at 1st and 2nd hidden layer respectively) trained on the preprocessing MNIST database. The best test error 0,8 %.

The preprocessing step with detector-descriptor scheme reduced test error of modified artificial neural network from 1.7 to 0.8 %.

The Yale Face Dataset. The Yale face databases were used in our experiments. The Yale face database1 contains 165 gray scale images of 15 individuals, each individual has 11 images. The images demonstrate variations in lighting condition, facial expression (normal, happy, sad, sleepy, surprised, and wink).

For the vector-based approaches, the image is represented as a 1 024-dimensional vector, while for the tensor-based approaches the image is represented as a (32×32)-dimensional matrix, or the second order tensor. The image set is then partitioned into the gallery and probe set with different numbers. For ease of representation, Gm/Pn means m images per person are randomly selected for training and the remaining n images are for testing.

First we obtain database using detector-descriptor scheme on the Yale database. For detector and descriptor we using the testing software provided by Mikolajczyk. The matching is carried out as follows. There were 1 000 2-layer modified artificial neural networks (168 hidden units: 165 and 3 units at 1st and 2nd hidden layer respectively) trained on the preprocessing Yale database. Each neuron on hidden layer was trained to identify the person. Table 2 summarizes the performance of algorithms compared at Yale data base [29]. For each Gp/Pq, the results average over 20 random splits and report the mean as well as the standard deviation.

Experimental results show that the detector-descriptor scheme as preprocessing step improve performance of modified artificial neural network significantly and detector-descriptor-modified artificial neural network scheme outperforms the ordinary subspace learning algorithms.

We have investigated detector-descriptor-modified artificial neural network scheme which fulfills the optimal complex and cross-validated model. Our analysis was based on object recognition and classification tasks. The modified

Table 1

Neural network results on MNIST

CLASSIFIER	Preprocessing	TEST ERROR RATE (%)	Reference
2-layer NN, 300 hidden units, mean square error	none	4.7	HLeCun et al. HH, 1998
2-layer NN, 300 HU, MSE, (distortions)	none	3.6	HLeCun et al. HH, 1998
2-layer NN, 300 HU	deskewing	1.6	HLeCun et al. HH, 1998
2-layer NN, 800 HU, cross-entropy (elastic distortions)	none	0.7	Simard et al., ICDAR, 2003
Convolutional net, cross-entropy [elastic distortions]	none	0.4	Simard et al., ICDAR, 2003
3-layer NN, 500+300 HU, softmax, cross entropy, weight decay	none	1.53	HHinton, unpublished, 2005
NN, 784-500-500-2000-30 + nearest neighbor, RBM + NCA training (no distortions)	none	1.00	Salakhutdinov and Hinton, AI-Stats, 2007
modified artificial neural network, 2 hidden layers, 33 HU	none	1.7	Engel, 2009
modified artificial neural network, 2 hidden layers, 33 HU	detector-descriptor scheme	0.8	Engel, 2009

neural network applied to solved tasks in different domains. Experimental results show that:

– the detector-descriptor-modified artificial neural network scheme effectively solves practical tasks of various subject fields and consistently outperforms the popular learning algorithms and advisable to gain extra prediction accuracy;

– the detector-descriptor scheme as preprocessing step improve performance of modified artificial neural network significantly;

– the detector-descriptor-modified artificial neural network scheme have done well to predict performance.

The bibliographic list

1. Lindeberg, T. Feature detection with automatic scale selection / T. Lindeberg // Intern. J. on Computer Vision. 1998. Vol. 30. № 2. P. 79–116.

2. Lowe, D. Distinctive image features from scale-invariant keypoints, cascade filtering approach / D. Lowe // Intern. J. on Computer Vision. 2004. Vol. 60. № 2. P. 91–110.

3. Mikolajczyk, K. An affine invariant interest point detector / K. Mikolajczyk, C. Schmid // Proc. of the Europ. Conf. on Computer Vision (ECCV). 2002. P. 128–142.

4. Ke, Y. PCA-SIFT: A more distinctive representation for local image descriptors / Y. Ke, R. Sukthankar // Proc. of the Conf. on Computer Vision and Pattern Recognition (CVPR). 2004. № 2. P. 506–513.

5. Tuytelaars, T. Wide baseline stereo based on local, affinely invariant regions // T. Tuytelaars, L. Van Gool // Proc. of the British Machine Vision Conf. (BMVC). 2000. P. 412–422.

6. Matas, J. Robust wide baseline stereo from maximally stable extremal regions / J. Matas, O. M. U. Chum, T. Pajdla // Proc. of BMVC. 2002. P. 384–393.

7. Mikolajczyk, K. A performance evaluation of local descriptors / K. Mikolajczyk, C. Schmid // Proc. of CVPR. 2003. Vol. 2. P. 257–263.

8. Mikolajczyk, K. A performance evaluation of local descriptors / K. Mikolajczyk, C. Schmid // IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI). 2005. Vol. 27. P. 1615–1630.

9. A comparison of affine region detectors / K. Mikolajczyk [et al.] // Intern. J. on Computer Vision. 2005. Vol. 65. № 1/2. P. 43–72.

10. Harris, C. A combined corner and edge detector / C. Harris, M. Stephens // Proc. of the Alvey Vision Conf. 1988. P. 147 – 151.

11. Mikolajczyk, K. Indexing based on scale invariant interest points / K. Mikolajczyk, C. Schmid // Proc. of the Intern. Conf. on Computer Vision (ICCV). 2001. Vol. 1. P. 525–531.

12. Lowe, D. Object recognition from local scale-invariant features / D. Lowe // Proc. of ICCV. 1999. Vol. 2. P. 1150–1157.

13. Kadir, T. Scale, saliency and image description / T. Kadir, M. Brady // Intern. J. on Computer Vision. 2001. Vol. 45. № 2. P. 83–105.

14. Jurie, F. Scale-invariant shape features for recognition of object categories / F. Jurie, C. Schmid // Proc. of the CVPR'04. 2004. Vol. 2. P. 90–96.

15. Mikolajczyk, K. Scale and affine invariant interest point detectors / K. Mikolajczyk, C. Schmid // Intern. J. on Computer Vision. 2004. Vol. 60. № 1. P. 63–86.

16. Florack, L. M. J. General intensity transformations and differential invariants / L. Florack [et al.] // J. of Math. Imaging and Vision. 1994. № 4. P. 171–187.

17. Mindru, F. Moment invariants for recognition under changing viewpoint and illumination / F. Mindru, T. Tuytelaars, L. Van Gool, T. Moons // Computer Vision and Image Understanding / 2004. Vol. 94. № 1–3. P. 3–27.

18. Baumberg, A. Reliable feature matching across widely separated views / A. Baumberg // Proc. of CVPR'00. 2000. Vol. 1 P. 774–781.

19. Schaffalitzky, F. Multi-view matching for unordered image sets, or "How do I organize my holiday snaps?" / F. Schaffalitzky, A. Zisserman // Proc. of ECCV. 2002. Vol. 1. P. 414–431.

20. Freeman, W. T. The design and use of steerable filters / W. T. Freeman, E. H. Adelson // IEEE Trans. PAMI. 1991. Vol. 13. P. 891–906.

21. Carneiro, G. Multi-scale phase-based local features / G. Carneiro, A. Jepson // Proc. of CVPR'03. 2003. Vol. 1. P. 736–743.

Table 2

Recognition accuracy on Yale (mean±std-dev %)

Method	G2/P9	G3/P8	G4/P7	G5/P6
Eigenface	46.0 ± 3.4	50.0 ± 3.5	55.7 ± 3.5	57.7 ± 3.8
Fisherface	45.7 ± 4.2	62.3 ± 4.5	73.0 ± 5.4	76.9 ± 3.2
2DLDA	43.4 ± 6.2	56.3 ± 4.7	63.5 ± 5.6	66.1 ± 4.8
S-LDA	57.6 ± 4.1	72.3 ± 4.4	77.8 ± 3.0	81.7 ± 3.2
Laplacianface	54.5 ± 5.2	67.2 ± 4.1	72.7 ± 4.2	75.8 ± 4.6
MFA	45.7 ± 4.2	62.3 ± 4.5	73.0 ± 5.4	76.9 ± 3.2
S-MFA	57.2 ± 4.3	71.2 ± 4.0	76.9 ± 3.1	81.1 ± 3.1
TensorPCA	49.4 ± 3.5	54.0 ± 3.0	57.8 ± 3.3	59.8 ± 3.9
S-LPP	57.9 ± 4.5	72.0 ± 4.0	76.0 ± 3.4	81.4 ± 2.9
S-NPE	57.5 ± 4.7	71.9±3.9	77.0 ± 3.4	80.9 ± 3.5
Pixel space	n/a	n/a	84.0 ± 1.5	n/a
Noushath et al. 2006	n/a	n/a	85.0 ± 1.5	n/a
Wang et al. 2007	n/a	n/a	99.0 ± 0.5	n/a
Modified artificial neural network	58.1 ± 3.5	72.1 ± 3.0	77.3 ± 3.4	81.7 ± 3.4
Detector-descriptor-modified artificial neural network	67.2 ± 2.8	79,8 ± 2,7	86.5 ± 2.9	91.6 ± 2.8

22. Se, S. Vision based modeling and localization for planetary exploration rovers / S. Se, H. Ng, P. Jasiobedzki, T. Moyung // Proc. of 55th Intern. Astronautical Cong. 2004. P. 1–11.

23. Viola, P. Rapid object detection using a boosted cascade of simple features / P. Viola, M. Jones // Proc. of CVPR'01. 2001. Vol. 1. P. 511–518.

24. Koenderink, J. The structure of images / J. Koenderink // Biol. Cybernetics. 1984. Vol. 50. P. 363–370.

25. Lindeberg, T. Discrete Scale-Space Theory and the Scale-Space Primal Sketch : PhD thesis / T. Lindeberg. Stockholm, 1991.

26. Lindeberg, T. Real-time scale selection in hybrid multi-scale representations / T. Lindeberg, L. Bretzner // Proc. Scale-Space'03. 2003. P. 148–163.

27. Brown, M. Invariant features from interest point groups / M. Brown, D. Lowe // Proc. of BMVC. 2002. P. 656–665.

28. Engel E. A. Modified artificial neural network for information processing with the selection of essential connections : PhD thesis / E. A. Engel. Krasnoyarsk, 2004.

29. Cai, D. Learning a Spatially Smooth Subspace for Face Recognition / D. Cai, X. He, Y. Hu, J. Han, T. Huang // Proc. of CVPR'07. 2007. P. 1–8.

Е. А. Энгель

ОБРАБОТКА ГРАФИЧЕСКОЙ ИНФОРМАЦИИ ИНТЕЛЛЕКТУАЛЬНЫМИ АЛГОРИТМАМИ

Основная проблема при построении системы распознавания образов – отыскание существенного набора свойств образа. Показано, что для распознания объектов с высокой изменчивостью формы, таких как рукописные цифры и лица, целесообразно использовать модифицированную нейронную сеть с предварительной предобработкой изображения объекта детектором и дескриптором, инвариантных к масштабированию и повороту. Эксперименты подтверждают эффективность схемы детектор-дескриптор-модифицированная нейронная сеть в реальных приложениях.

Ключевые слова: распознавание образов, схема детектор-дескриптор-модифицированная нейронная сеть.

А. А. Городилов

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДИНАМИЧЕСКИХ СТРУКТУР ДАННЫХ АВТОМАТИЗИРОВАННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

Рассмотрены вопросы хранения информации в динамических структурах данных как наиболее эффективном и адаптивном к изменениям самих структур способе их организации. Рассмотрена математическая модель, описывающая структуру и взаимосвязи данных предметной области, используя динамические структуры данных.

Ключевые слова: математические модели, информационные системы, разработка, динамические структуры данных.

В настоящее время происходит массовое внедрение автоматизированных информационных систем (АИС) обработки данных в различных организациях, что связано с ужесточением требований к оперативности и качеству обрабатываемой информации и увеличением объемов информации, которую им нужно обрабатывать. Также в настоящее время происходят различные изменения в бизнес-процессах организаций, законодательстве, документообороте и др.

При поддержке АИС, разработанных согласно стандартным принципам и нацеленным на удовлетворение текущих потребностей автоматизации, в подобных условиях возникает множество проблем, таких как необходимость постоянной адаптации АИС к изменяющимся условиям, высокая стоимость поддержки АИС (40…100 % от первоначальной стоимости в год), потеря собранных данных в связи с изменениями документов.

Эти проблемы возможно решить, создавая АИС, использующие динамические структуры данных (АИСДСД) [1], характеризующиеся возможностью гибкой перестройки структуры без потерь существующей информации.

Динамические структуры данных характеризуются отсутствием физической смежности элементов в памяти, а их логическая структура не связана с последовательностью размещения на физическом уровне, непостоянством и непредсказуемостью размера (числа элементов) структуры в процессе ее обработки. Для установления связи между элементами структуры используются ука-