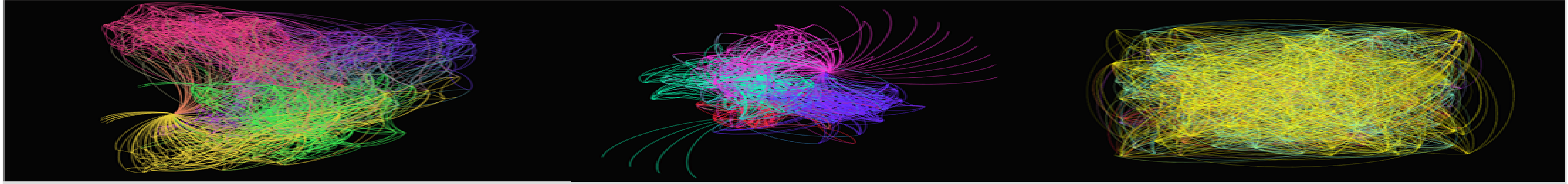


Workshop on Clustering and Search techniques in large scale networks - 2014

National Research University
Higher School of Economics
Nizhny Novgorod

Piecewise-regular object recognition in real-time applications

Presenter: **Andrey V. Savchenko**
National Research University Higher School of Economics
Department of Information Systems and Technology,
PhD, Associate Professor
Email: avsavchenko@hse.ru

Outline

1. Overview (classification of classification procedures)
2. Probabilistic Neural Networks (PNN)
3. Approximate nearest neighbor search
4. Experimental results (1). Face recognition.
5. Experimental results (2). Speech recognition in voice control applications
6. Conclusion

Introduction

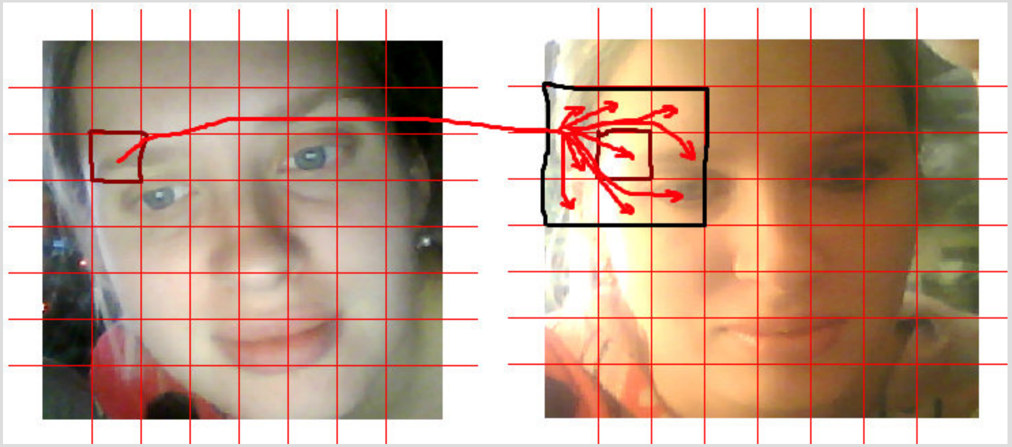
Problem. Let the database of *R>1* model objects X_r is given. Each *r*-th model is put in correspondence with label *c(r)*. Number of classes $C \le R$.

Query object

- Feature vector (traditional classification methods)
- Simple sample (group) of feature vectors (group choice classifiers)
- Sequence of segments (groups) of feature vectors (composite object classification)

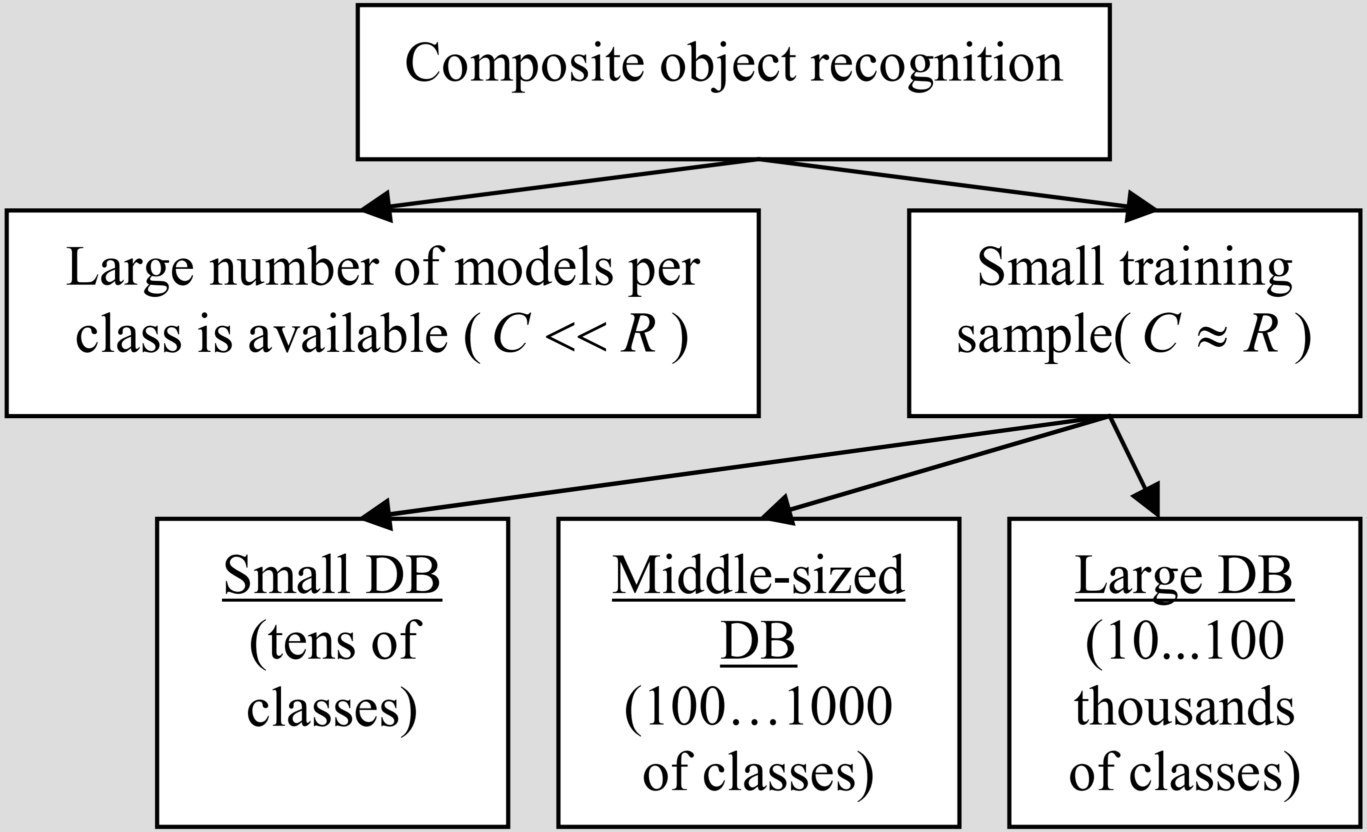
Composite (piecewise-regular) objects:
-images (HOG, SIFT, blocks in JPEG/MPEG)
-speech (phonemes/triphones).

Query object *X* and each model X_r are considered as sequences of, respectively, *K* and K_r homogeneous (regular) segments (phones). These segments are relatively independent (features of different segments inside one word may have nothing in common as they correspond to distinct phones).



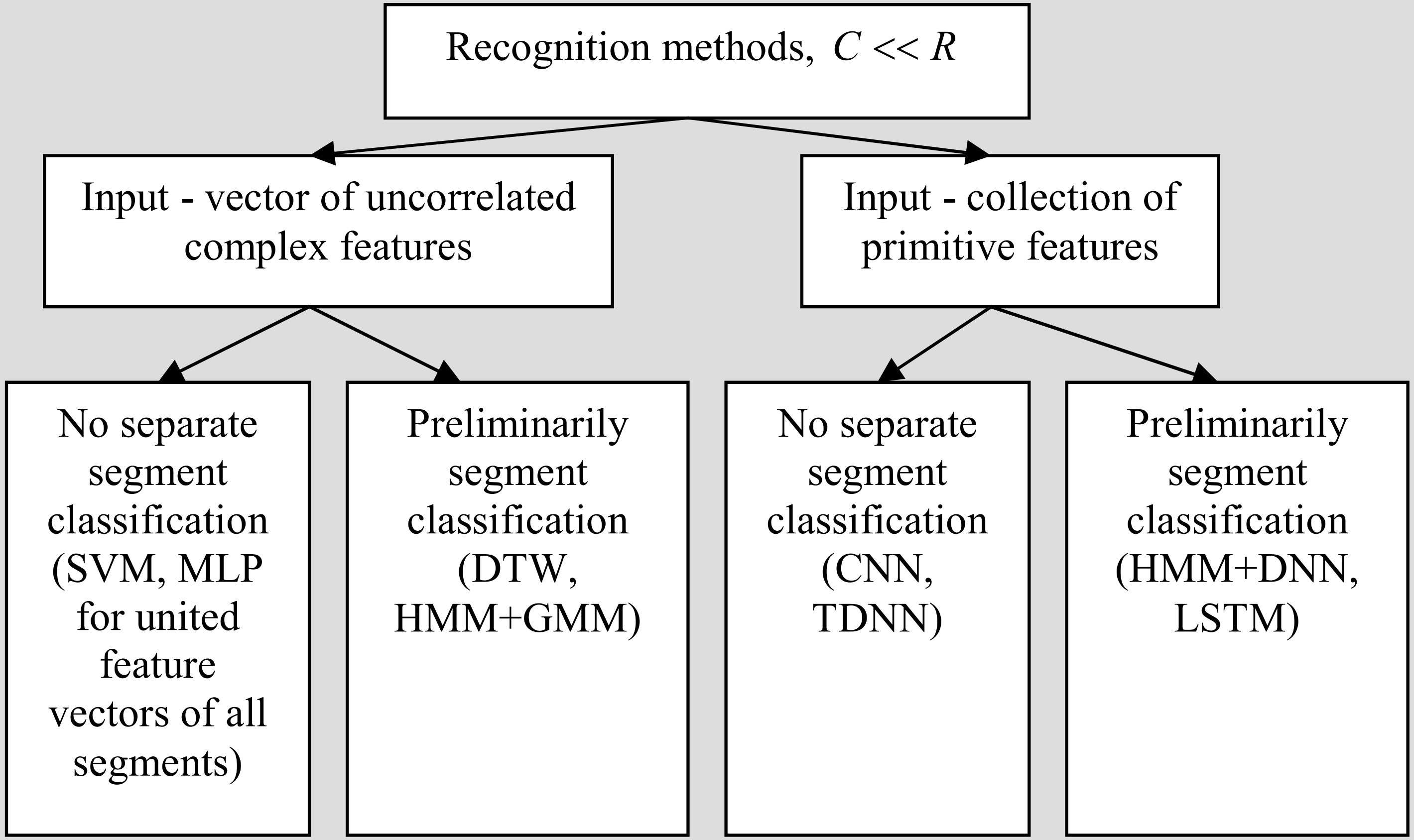
Composite objects’ recognition methods

Classification of composite object recognition systems in dependence on the available model database



The decision is made in favor of the closest model in terms of the summary distance for all K segments. Each segment is recognized with pointwise or group choice classifiers. As the number of extracted segments is usually high, computing efficiency of algorithms of composite object recognition is rather low

Large number of models per class (*C*<<*R*)

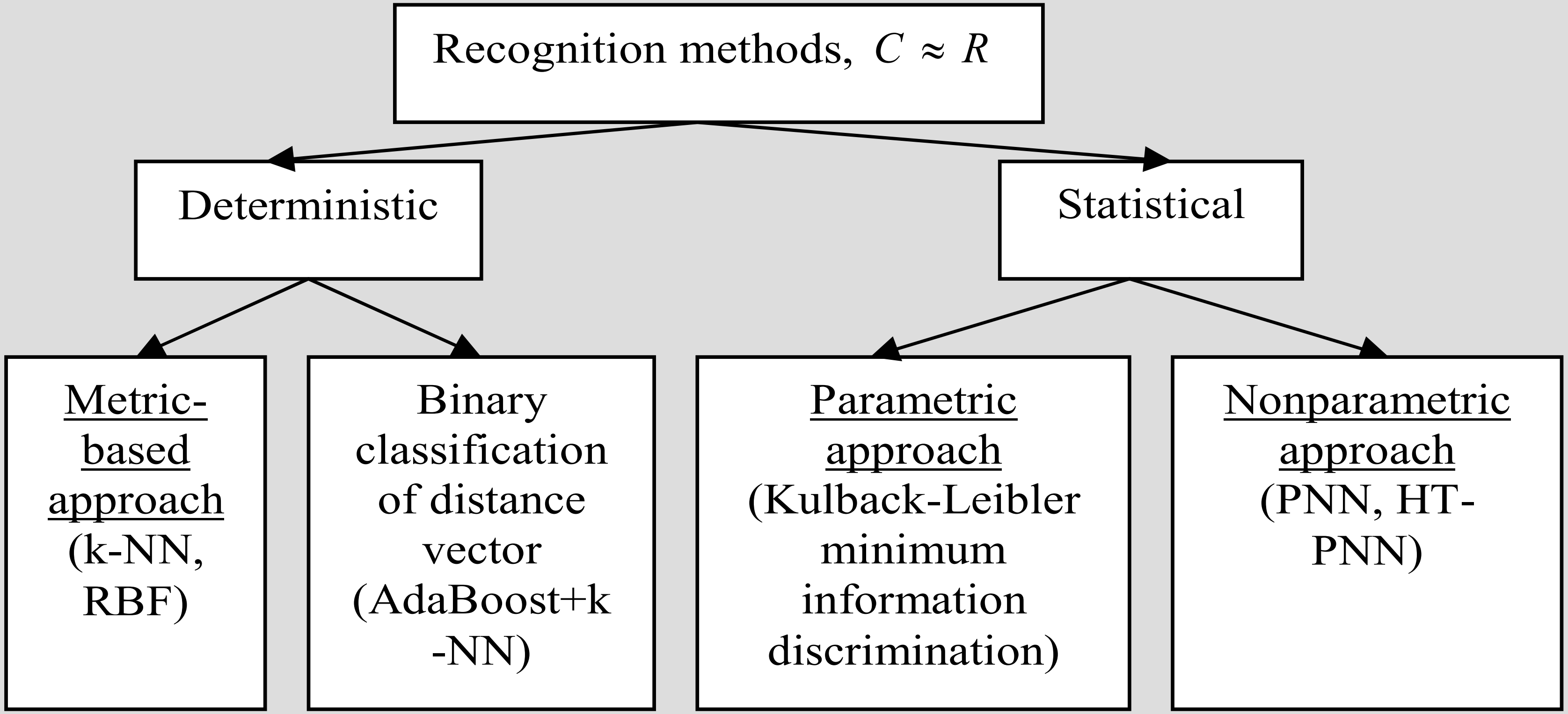


Nowadays most popular is the second approach implemented in deep neural networks (DNN).

1. Image recognition. Excellent results were shown with Multi-Column GPU Max-Pooling Convolutional Neural Network (MC-GPU-MPCNN).
2. Speech recognition.

- simple FFT-based features allow the DNN to achieve better accuracy than the conventional MFCC.
- recurrent LSTM (Long Short-Term Memory) trained with the Connectionist Temporal Classification outperforms state-of-the-art for TIMIT phonetic transcription

Small Training Sample



In *parametric* approach the type of distribution is fixed. Usually Gaussian approximation (in LDA/QDA) or polynomial distribution (histograms) for discrete features are used. It is possible to show that this approach is equivalent to the Kullback-Leibler minimum information discrimination principle if the segments are considered as a simple random sample of independent identically distributed primitive features

Practically all recognition methods for this case are implemented in k-NN (nearest neighbor) rules

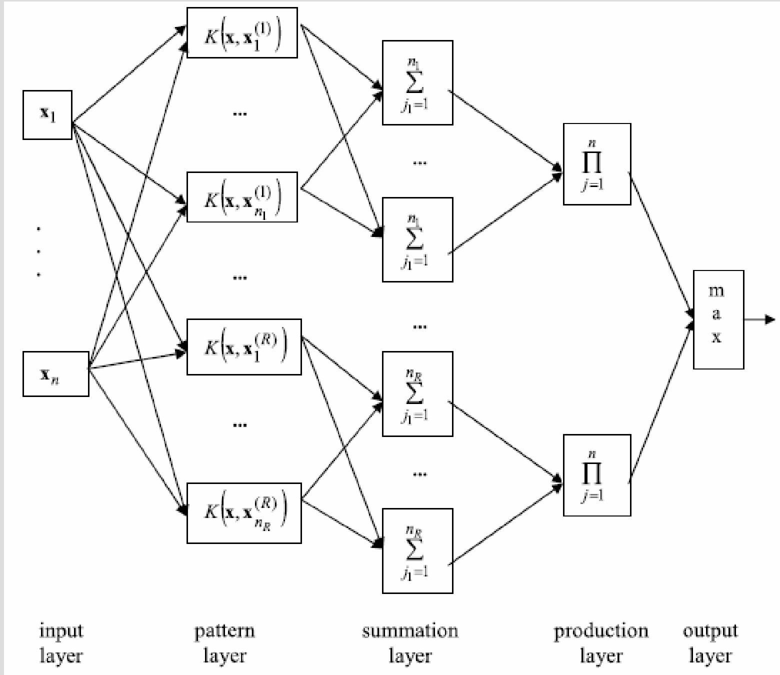
Probabilistic Neural Network (PNN) in group-choice classification

Classification task is reduced to the testing of **simple hypothesis**. In case of equal prior probabilities the Bayesian rule is as follows:

$$f\left(X(k)\middle|W_r(k_1)\right) \to \max_{r\in\{1,\ldots,R\}}$$

Instead of unknown class distributions let's use the Gaussian Parzen *kernel*.

Thus, for group-choice classification of the segment and naïve assumption of features independence inside each segment *the generalized PNN is used*



$$f\left(X(k)\middle|W_r(k_1)\right) = \frac{1}{\left(n_r(k_1)\right)^{n(k)}} \prod_{j=1}^{n(k)} \sum_{j_r=1}^{n_r(k_1)} K\left(\mathbf{x}_j(k), \mathbf{x}_{j_r}^{(r)}(k_1)\right)$$

Production layer is added to the traditional PNN structure

Final classifier with assumption of segments independence:

$$\sum_{k=1}^{K} \max_{k_1 \in N_r(k)} \frac{1}{\left(n_r(k_1)\right)^{n(k)}} \prod_{j=1}^{n(k)} \sum_{j_r=1}^{n_r(k_1)} K\left(\mathbf{x}_j(k), \mathbf{x}_{j_r}^{(r)}(k_1)\right) \to \max_{r=\overline{1,R}}$$

Unfortunately, if the distribution estimates are used instead of unknown distributions, such approach is **not optimal**. It is necessary to check **complex** hypothesis

Homogeneity-Testing PNN (HT-PNN)

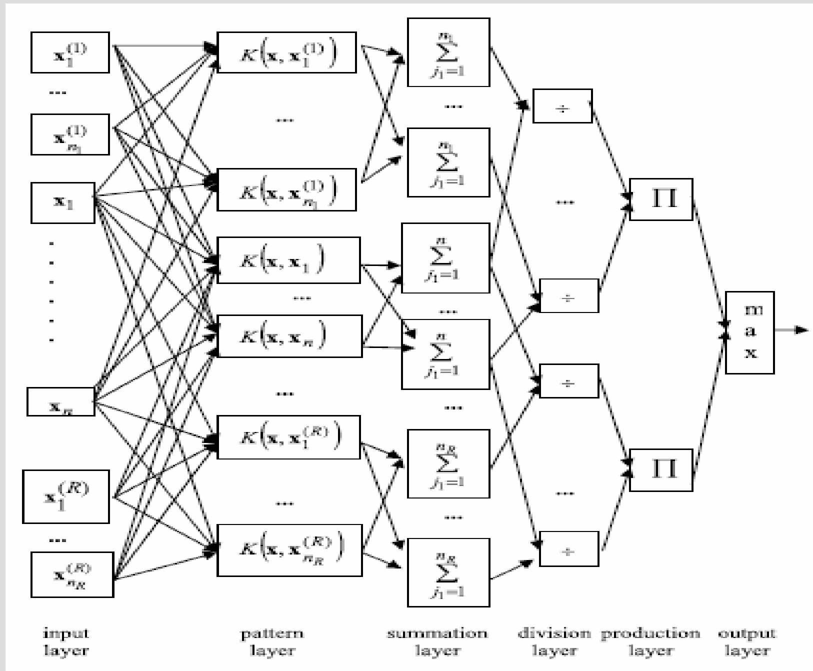
Idea - it is necessary to check **complex** hypothesis of features samples **homogeneity**. The following criterion is known to be asymptotically minimax:

$$\sup_{f^{(1)}_{k;k_1},\dots,f^{(R)}_{k;k_1}} f\left(\{X(k),X_1(k_1),\dots,X_R(k_1)\} \middle| W^{(r)}_{k;k_1}\right) \to \max_{r\in\{1,\dots,R\}} \quad\Rightarrow\quad \frac{\sup\limits_{f^{(r)}_{k;k_1}} f^{(r)}_{k;k_1}\left(X(k) \middle| W^{(r)}_{k;k_1}\right) f^{(r)}_{k;k_1}\left(X_r(k_1) \middle| W^{(r)}_{k;k_1}\right)}{\sup\limits_{f^{(r)}_{k;k_1}} f^{(r)}_{k;k_1}\left(X_r(k_1)\right)} \to \max_{r\in\{1,\dots,R\}}$$

Distribution of hypothesis $W^{(r)}_{k;k_1}$ is estimated by the united sample $\{X(k), X_r(k_1)\}$, as it is done in **Lehmann-Rosenblatt** test.

HT-*PNN for piecewise-regular object recognition*

$$\sum_{k=1}^{K} \max_{k_1 \in N_r(k)} \frac{n(k)^{n(k)} \cdot (n_r(k_1))^{n_r(k_1)}}{(n(k)+n_r(k_1))^{n(k)+n_r(k_1)}} \times$$
$$\times \prod_{j=1}^{n(k)} \left(1 + \frac{\sum\limits_{j_r=1}^{n_r(k_1)} K\left(\mathbf{x}_j(k), \mathbf{x}^{(r)}_{j_r}(k_1)\right)}{\sum\limits_{j_1=1}^{n(k)} K\left(\mathbf{x}_j(k), \mathbf{x}_{j_1}(k_1)\right)} \right) \times$$
$$\times \prod_{j_r=1}^{n_r(k_1)} \left(1 + \frac{\sum\limits_{j_1=1}^{n(k)} K\left(\mathbf{x}^{(r)}_{j_r}(k_1), \mathbf{x}_{j_1}(k)\right)}{\sum\limits_{j_{r;1}=1}^{n_r(k_1)} K\left(\mathbf{x}^{(r)}_{j_r}(k_1), \mathbf{x}^{(r)}_{j_{r;1}}(k_1)\right)} \right) \to \max_{r=\overline{1,R}}$$



If $n_r(k_1) \to \infty$, then the PNN and the HT-PNN are equivalent

Discrete features

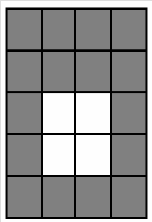
Computing efficiency of the PNN and the HT-PNN $O\left(M\cdot\sum_{r=1}^{R}\sum_{k=1}^{K}\sum_{k_1\in N_r(k)} n(k)\cdot n_r(k_1)\right)$

Practical implementation for hundreds of features and *R=100* **is impossible**

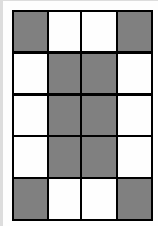
If there are only *N* feature values, segments *X(k)*, $X_r(k_1)$ are described with their histograms $\{w_i(k)\},\{\theta_{r;i}(k_1)\}, i=\overline{1,N}$

<table>
<tr><td>PNN</td><td>Equivalent to the Kullback-Leibler divergence, if smoothing parameter σ→0</td></tr>
<tr><td colspan="2">$$\rho_{PNN}(X,X_r)=\frac{1}{Kn}\sum_{k=1}^{K}\min_{k_1\in N_r(k)}\sum_{i=1}^{N} w_i(k)\ln\frac{w_{K;i}(k)}{\theta^{(r)}_{K;i}(k_1)}\to\min_{r\in\{1,\dots,R\}}$$
$$\theta^{(r)}_{K;i}(k_1)=\sum_{j=1}^{N}K_{ij}\theta^{(r)}_{j}(k_1)$$
$$n=\sum_{k=1}^{K}n(k)/K$$
$$w_{K;i}(k)=\sum_{j=1}^{N}K_{ij}w_j(k)$$</td></tr>
<tr><td>HT-PNN</td><td>Generalization of the Jensen-Shannon divergence</td></tr>
<tr><td colspan="2">$$\rho_{HT-PNN}(X,X_r)=\frac{1}{Kn}\sum_{k=1}^{K}\min_{k_1\in N_r(k)}\sum_{i=1}^{N}\left(n(k)w_i(k)\ln\frac{w_{K;i}(k)}{\tilde{\theta}^{(r)}_{\Sigma;i}(k;k_1)}+n_r(k_1)\theta^{(r)}_i(k_1)\ln\frac{\theta^{(r)}_{K;i}(k_1)}{\tilde{\theta}^{(r)}_{\Sigma;i}(k;k_1)}\right)\to\min_{r\in\{1,\dots,R\}}$$
$$\tilde{\theta}^{(r)}_{\Sigma;i}(k;k_1)=(n_r(k_1)\cdot\theta^{(r)}_{K;i}(k_1)+n(k_1)\cdot w_{K;i}(k))/\left(n_r(k_1)+n(k_1)\right)$$</td></tr>
<tr><td>Approximate HT-PNN (A-HT-PNN)</td><td>Generalization of chi-square distance</td></tr>
<tr><td colspan="2">$$\rho_{A-HT-PNN}(X,X_r)=\frac{1}{2Kn}\sum_{k=1}^{K}\min_{k_1\in N_r(k)}\sum_{i=1}^{N}\left(n(k)w_i(k)\left(\frac{w_{K;i}(k)}{\tilde{\theta}^{(r)}_{\Sigma;i}(k;k_1)}-\frac{\tilde{\theta}^{(r)}_{\Sigma;i}(k;k_1)}{w_{K;i}(k)}\right)+n_r(k_1)\theta^{(r)}_i(k_1)\left(\frac{\theta^{(r)}_{K;i}(k_1)}{\tilde{\theta}^{(r)}_{\Sigma;i}(k;k_1)}-\frac{\tilde{\theta}^{(r)}_{\Sigma;i}(k;k_1)}{\theta^{(r)}_{K;i}(k_1)}\right)\right)\to\min_{r\in\{1,\dots,R\}}$$</td></tr>
</table>

Example

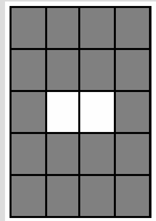


Class 1
(0,8; 0,2)

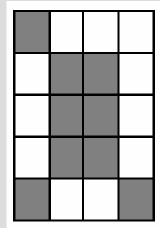


Class 2
(0,5; 0,5)

But models can contain mistakes!

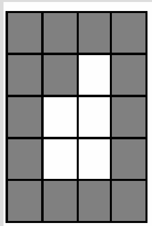


Model 1
(0,9; 0,1)



Model 2
(0,55; 0,45)

Query object
(0,75; 0,25) from class 1



PNN_dist(test, model1)=0,092 PNN_dist(test, model2)=0,086

HT-PNN compares:

a) query object *X* and model 1 with distribution (0,825; 0,175):

HT-PNN_dist(test, model1)=0,040

б) query object *X* and model 2 with distribution (0,65; 0,35):

HT-PNN_dist(test, model2)=0,044

For binary images (size 4x5) in 18 cases the HT-PNN for 2 classes gives different results than the PNN. In ALL these cases HT-PNN makes decision in favor of the class which model has lower number of pixels different with the query object

Analysis of efficiency

Computing efficiency is in n^2M/N times higher: $$O\left(2N \cdot \sum_{r=1}^{R} \sum_{k=1}^{K} |N_r(k)|\right)$$

It is known the relation of the likelihood ratio with the noncentral chi-square (see Kullback, 1997). If *KN* is high, distance in the HT-PNN between object of class *v* and model X_r is asymptotically distributed as

$$N\left(\rho_{HT-PNN}(X_\nu, X_r) + \frac{N-1}{n}; \frac{\sqrt{4nK \cdot \rho_{HT-PNN}(X_\nu, X_r) + 2K \cdot (N-1)}}{nK}\right)$$

Hence the error rate for the objects of *v*–th lass is estimated from the distribution of minimum independent normal random variables

$$\alpha_\nu = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp(-t^2/2) \prod_{i \in \{1,\ldots,\nu-1,\nu+1,\ldots,R\}} \left(\frac{1}{2} - \frac{1}{2}\Phi(\tilde{t}(t,\nu,i))\right) dt$$

$\Phi(\tilde{t})$ - cdf of the normal distribution

$$\tilde{t}(t,\nu,i) = \frac{t \cdot \sqrt{2K \cdot (N-1)} - K \cdot (N-1) \cdot \rho_{\nu,i}}{\sqrt{4nK \cdot \rho_{\nu,i} + 2K \cdot (N-1)}} \qquad \rho_{\nu,i} = \rho_{PNNH}(X_\nu, X_i)$$

The more the distance between X_v and other models, the higher is the recognition rate for objects from the *v*-th class.

Small Training Sample (2). Large database

k-NN rule requires the brute force of the whole database

Small training sample $C \approx R$

- Small database (tens of classes)
- Medium-sized DB. Problems with accuracy and computational speed *automatic* real-time recognition
- Very-large DB. *Automated* content-based object retrieval (approximate k-NN)

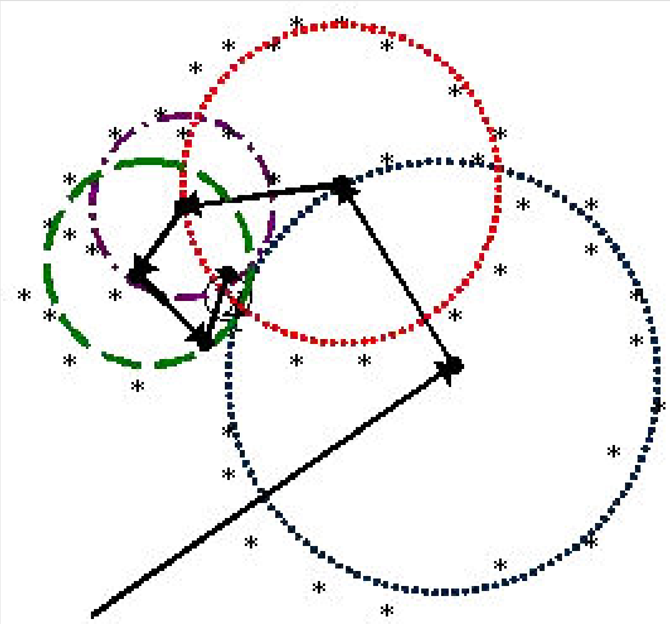
Solutions:

- Modern hardware
- Parallel computing
- Simplification of similarity measure or its parameters.
- Approximate NN methods looks for the model which is the nearest neighbor with high probability. Example:

$$\rho(X, X_\nu) < \rho_0 = const$$

Most methods are implemented for Minkowski distances (e.g. AESA – triangle tree). Problems with measures which do not satisfy metric properties (symmetry and triangle inequality).

Medium-sized databases. Directed Enumeration Method (DEM)



Idea: on each step the next model is selected to maximize likelihood of the previously calculated distances

$$r_{k+1} = \underset{\nu \in \{1,\dots,R\} - \{r_1,\dots,r_k\}}{\arg\max} \prod_{i=1}^{k} f\left(\rho_{HT-PNN}\left(X, X_{r_i}\right) \middle| W_\nu\right)$$

Based on the asymptotic properties of the HT-PNN, this rule is equivalent to

$$r_{k+1} = \underset{\mu \in \{1,\dots,R\} - \{r_1,\dots,r_k\}}{\arg\min} \sum_{i=1}^{k} \varphi_\mu(r_i) \qquad \varphi_\mu(r_i) \approx \frac{\left(\rho_{HT-PNN}\left(X, X_{r_i}\right) - \rho_{\mu,r_i}\right)^2}{\rho_{\mu,r_i}}$$

Initialization: r_1 is chosen to maximize average probability to obtain correct decision of $k=2$-th step

$$r_1 = \underset{\mu \in \{1,\dots,R\}}{\arg\max} \sum_{\nu=1}^{R} \prod_{r=1}^{R} \left(\frac{1}{2} + \Phi\left(\frac{\sqrt{nK}}{2} \left| \sqrt{\rho_{r,\mu}} - \sqrt{\rho_{\nu,\mu}} \right| \right) \right)$$

Termination: the model v is the solution if

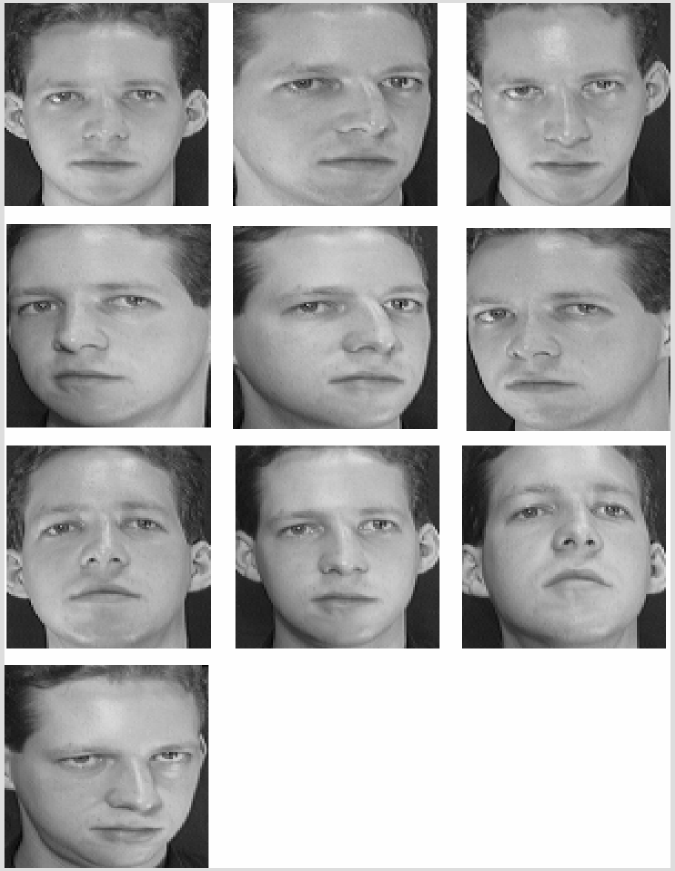
$$\rho(X, X_\nu) < \rho_0 = const$$

1. In asymptotic (n, $n_r \to \infty$) the number of distance calculations in the DEM is constant (doe not depend on the DB size R)
2. The DEM is the optimal greedy algorithm for the HT-PNN.

Experimental results (1). Face recognition

FERET (2720 face-to-face images of 994 persons)

AT&T (400 images of 40 persons)



Testing: 20-time repeated random subsampling cross-validation.
Random noise from the range $[-x_{noise}; x_{noise}]$ is added to every pixel of the test image

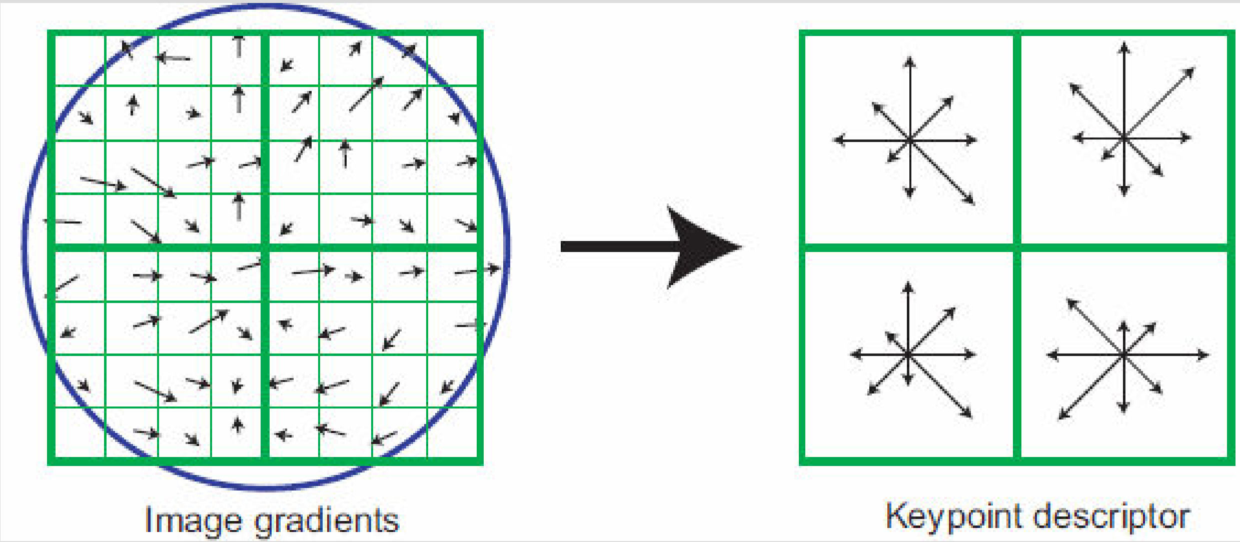
Histograms of Oriented Gradients (HOG)

Proposed in (Dalal and Triggs, 2005)

Image descriptors:

1. Local (SIFT, SURF, etc.):
   a) Keypoint extraction
   b) Descriptor extraction
2. Global (color histograms, HOG)
   a) Object detection
   b) Descriptor extraction

Gradient orientation histogram (from (Lowe, 2004)):

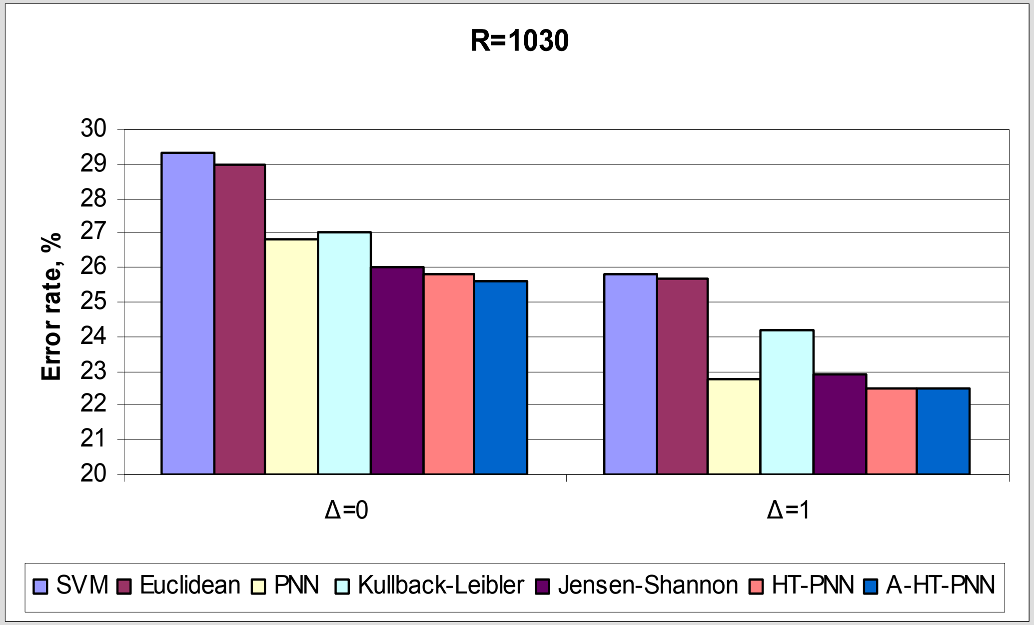


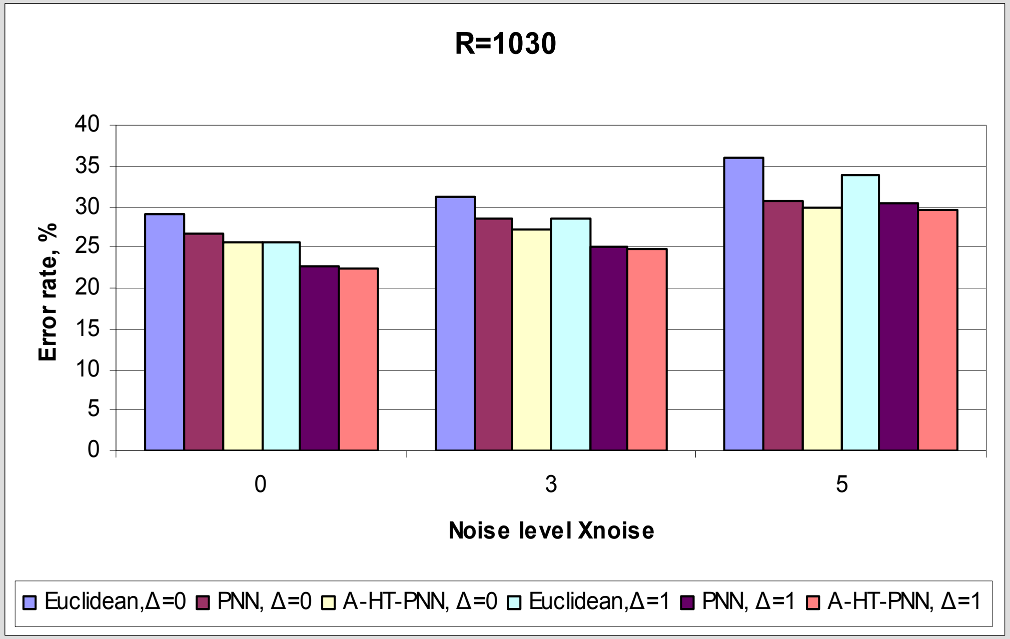
Parameters: 10x10 grid.

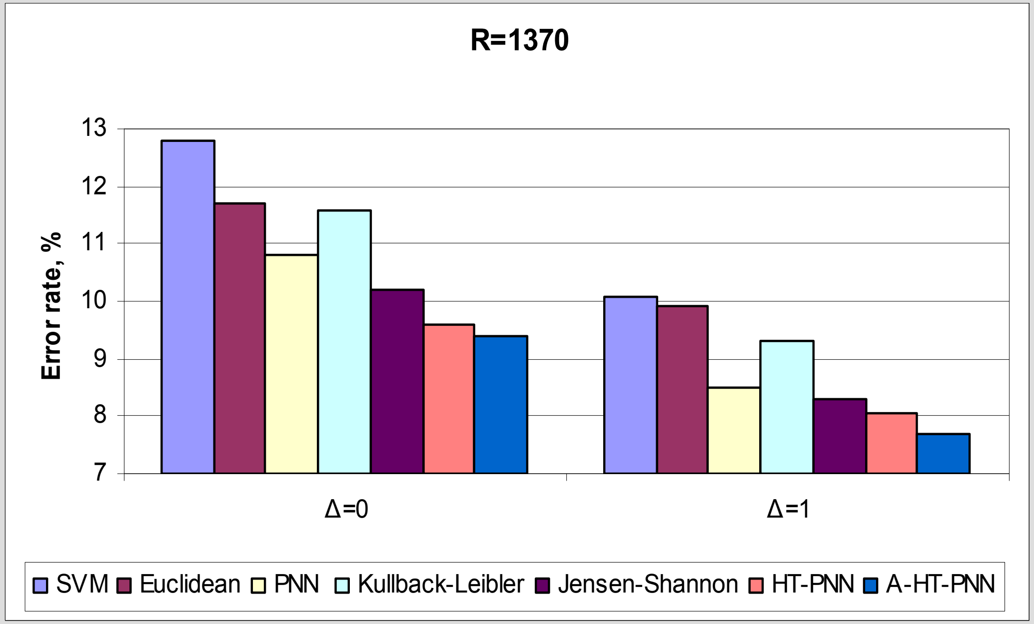
Criterion:

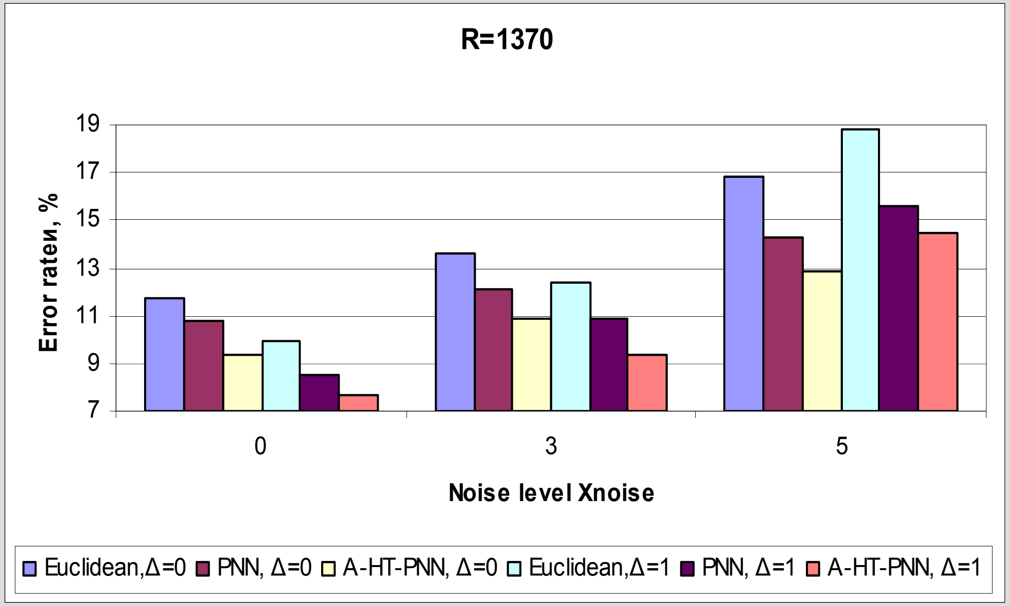
$$\frac{K^{(1)}K^{(2)}}{UV}\sum_{k_1=1}^{K^{(1)}}\sum_{k_2=1}^{K^{(2)}}\min_{\substack{|\Delta_1|\le\Delta,\\|\Delta_2|\le\Delta}}\rho\left(H_r\left(k_1+\Delta_1,k_2+\Delta_2\right),H\left(k_1,k_2\right)\right)\to\min_{r\in\{1,\dots,R\}}$$

Experimental results (1). FERET dataset

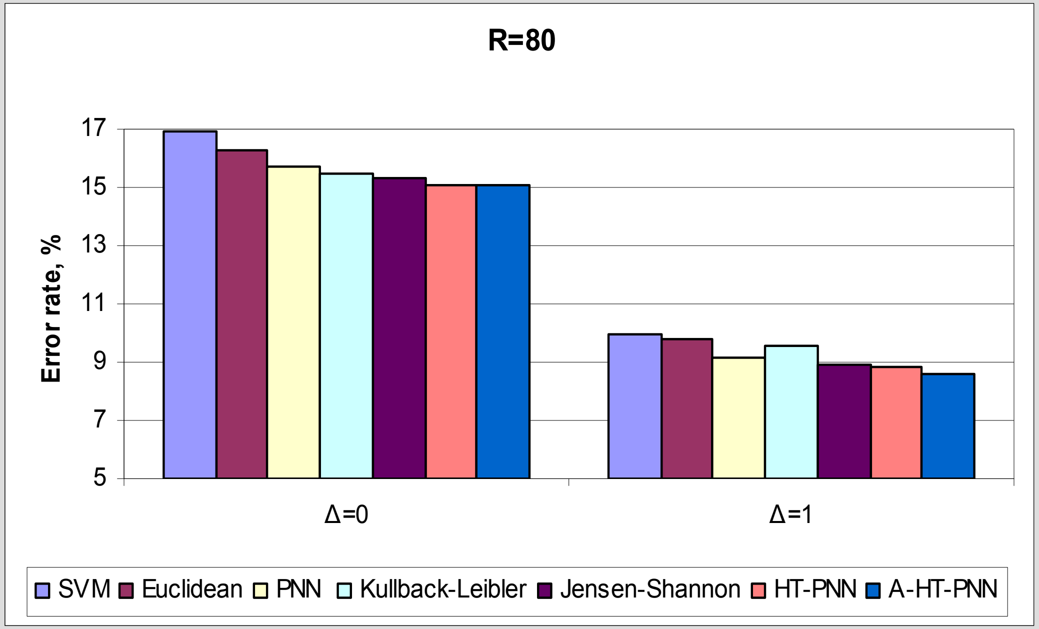


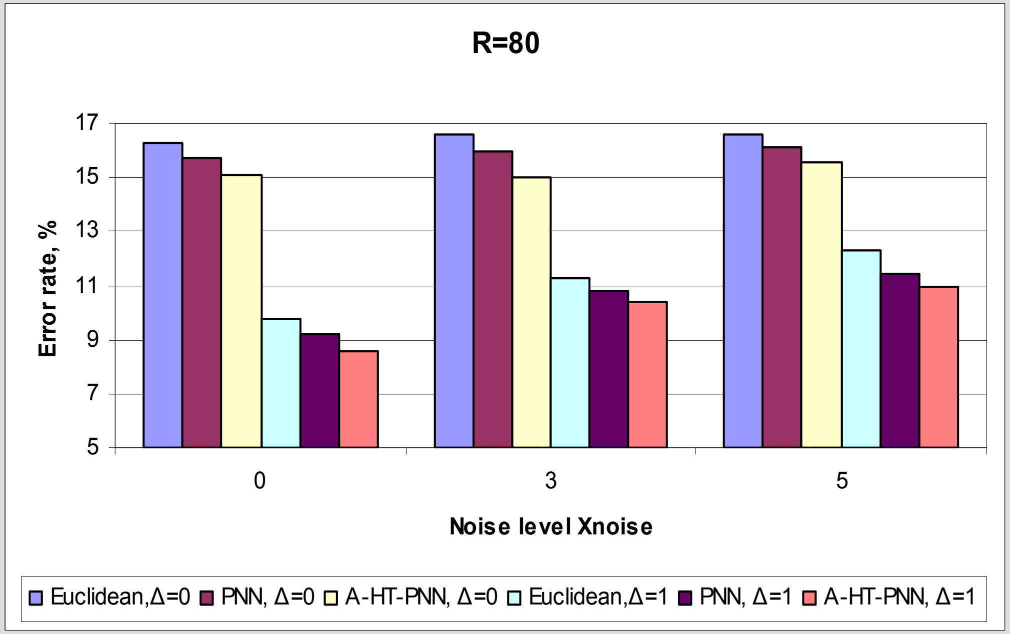


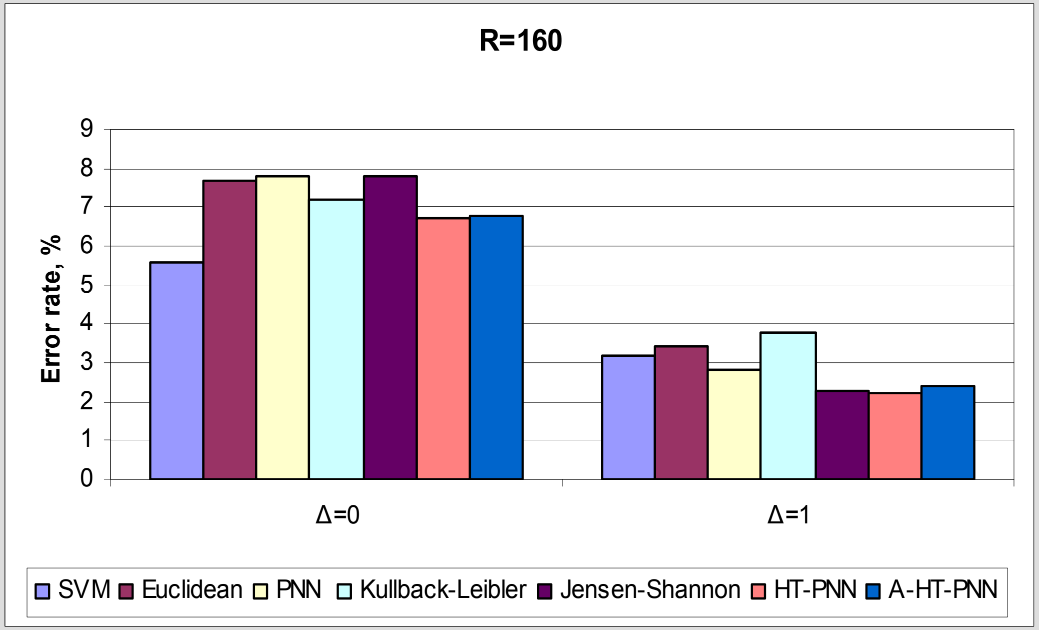


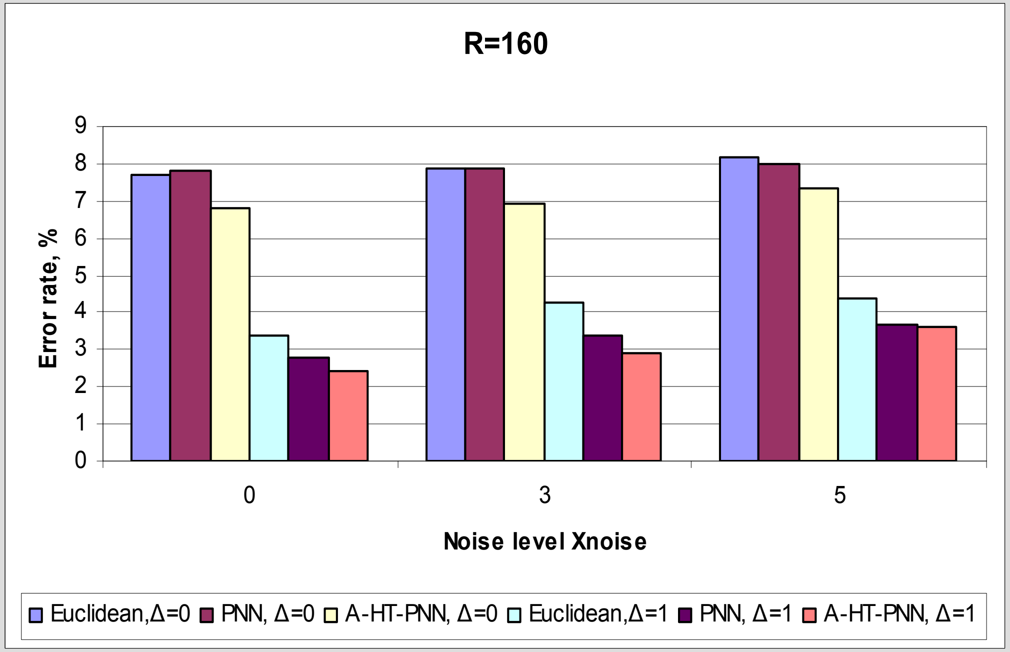


Experimental results (1). AT&T dataset



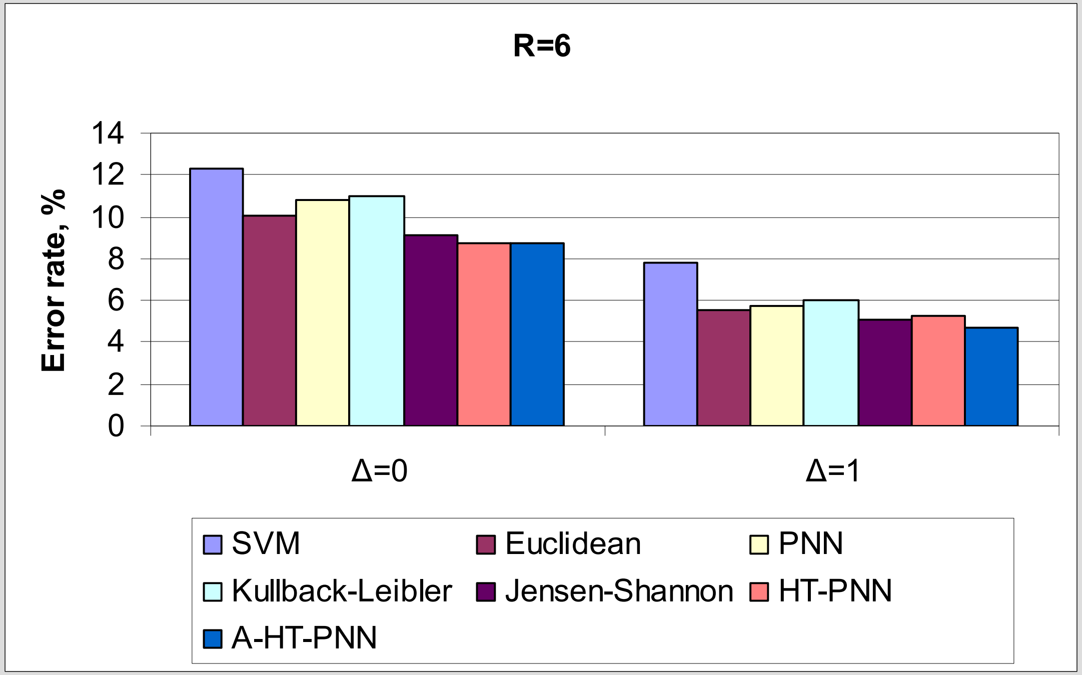


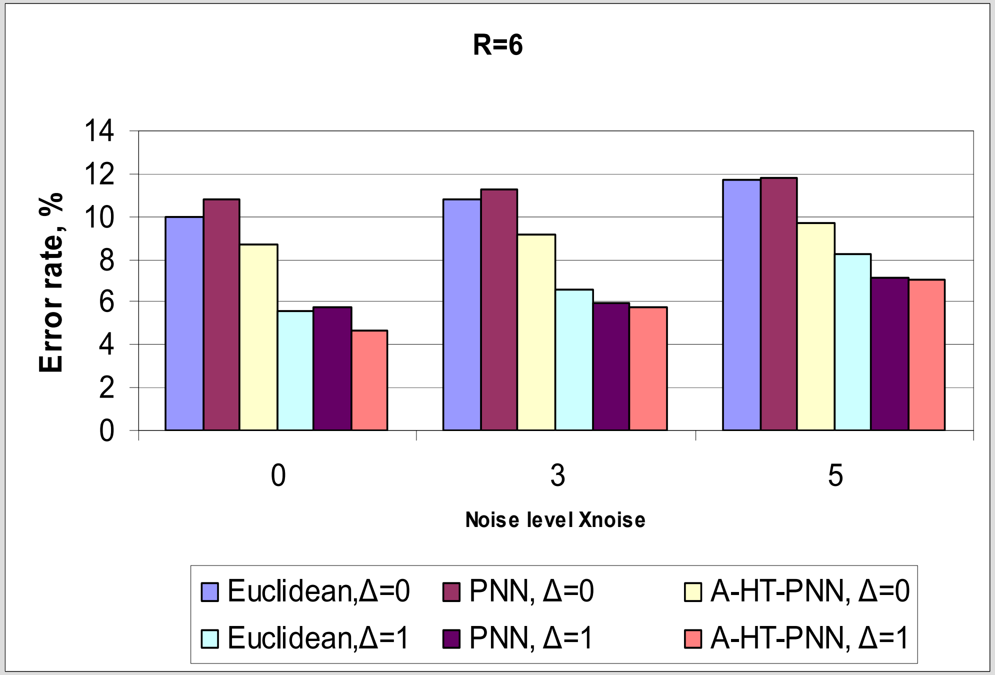




Experimental results (1). Face recognition. Real data

120 photos of 6 students (4 male, 2 female)



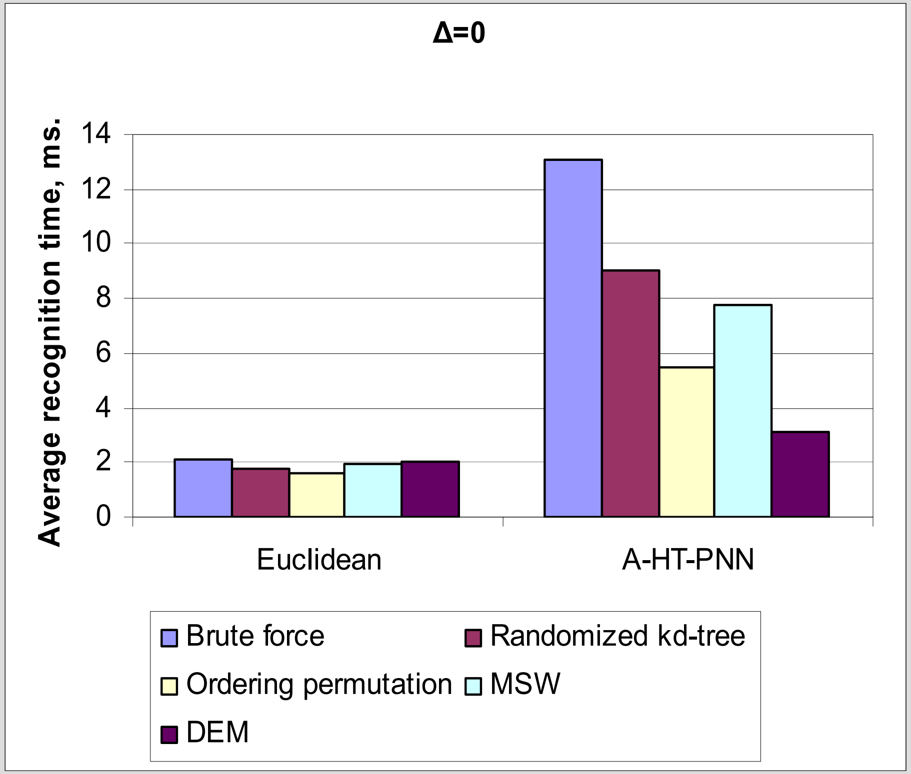


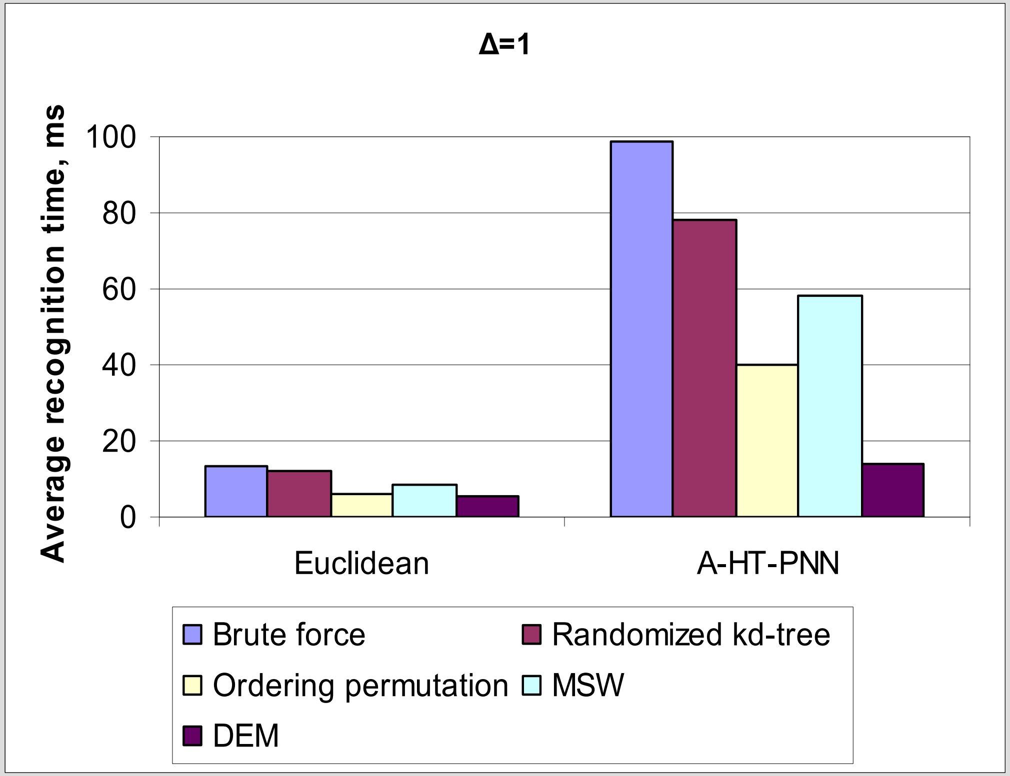
Experimental results (2). DEM

The DEM is compared with the following approximate NN methods

1. **Randomized kd-tree**. Silpa-Anan C., Hartley R. Optimised KD-trees for fast image descriptor matching // IEEE Int. Conf. on CVPR. 2008.
2. **Ordering permutation**. Gonzalez E.C., Figueroa K., Navarro G. Effective Proximity Retrieval by Ordering Permutations // IEEE Trans. on PAMI. 2008
3. **Metrized Small World (MSW).** Malkov Y. et al. Approximate nearest neighbor algorithm based on navigable small world graphs // Information Systems. 2014

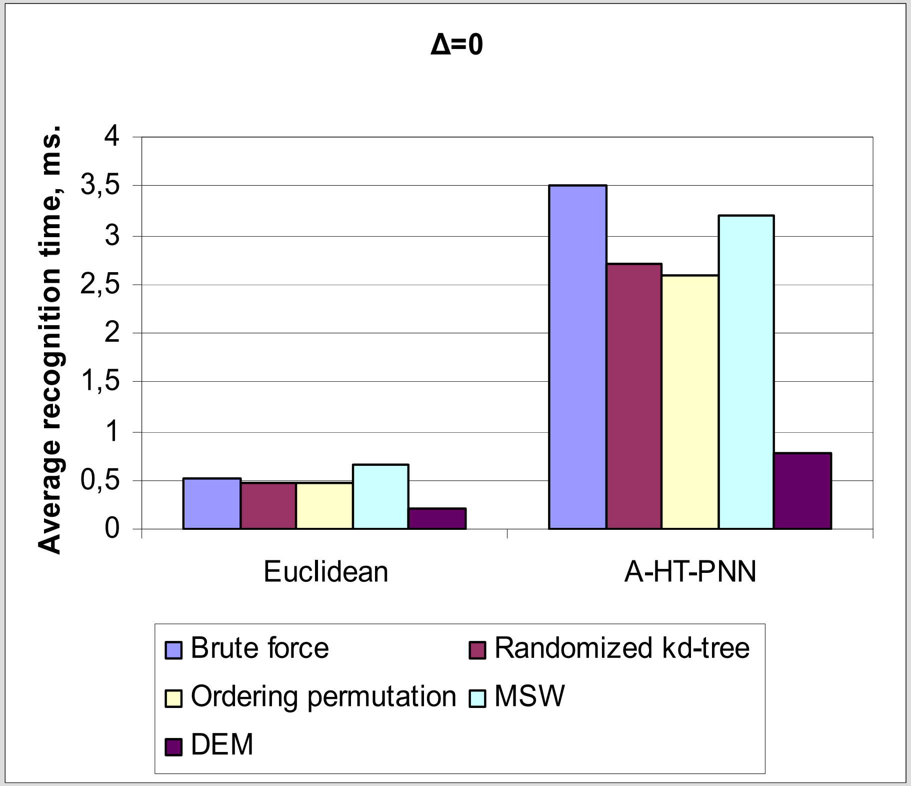
Average recognition time (ms.), FEREТ, *R=1400*, laptop Core i7-2630QM, 2 GHz, RAM 6 Gb, Visual C++ 2013 Express, x64.

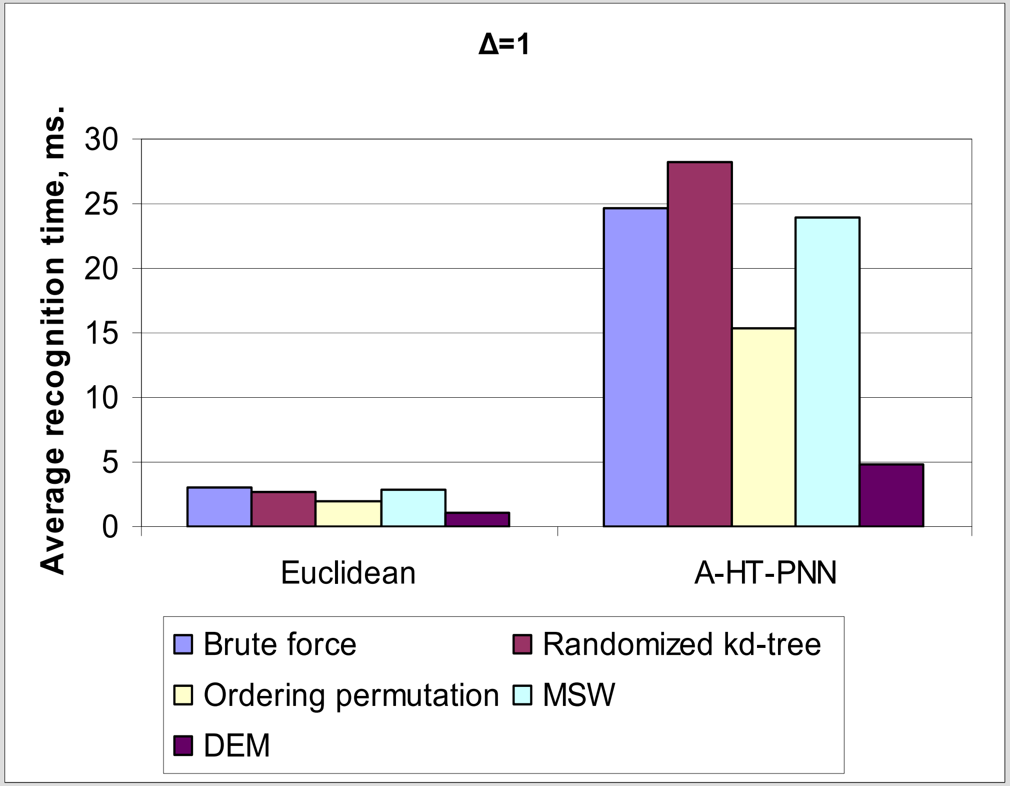




Experimental results (2). Parallel DEM

Average recognition time (ms.), T=8 parallel threads





1. Speed of conventional approximate NN methods is usually unsatisfactory;
2. Parallel DEM is characterized with the best time: 2,5-7 times higher in comparison with single threaded 1-NN and 2,5-5 times higher in comparison with 8-threaded 1-NN with comparable accuracy;
3. Increase of the computational speed for the DEM is higher for more complex similarity measure

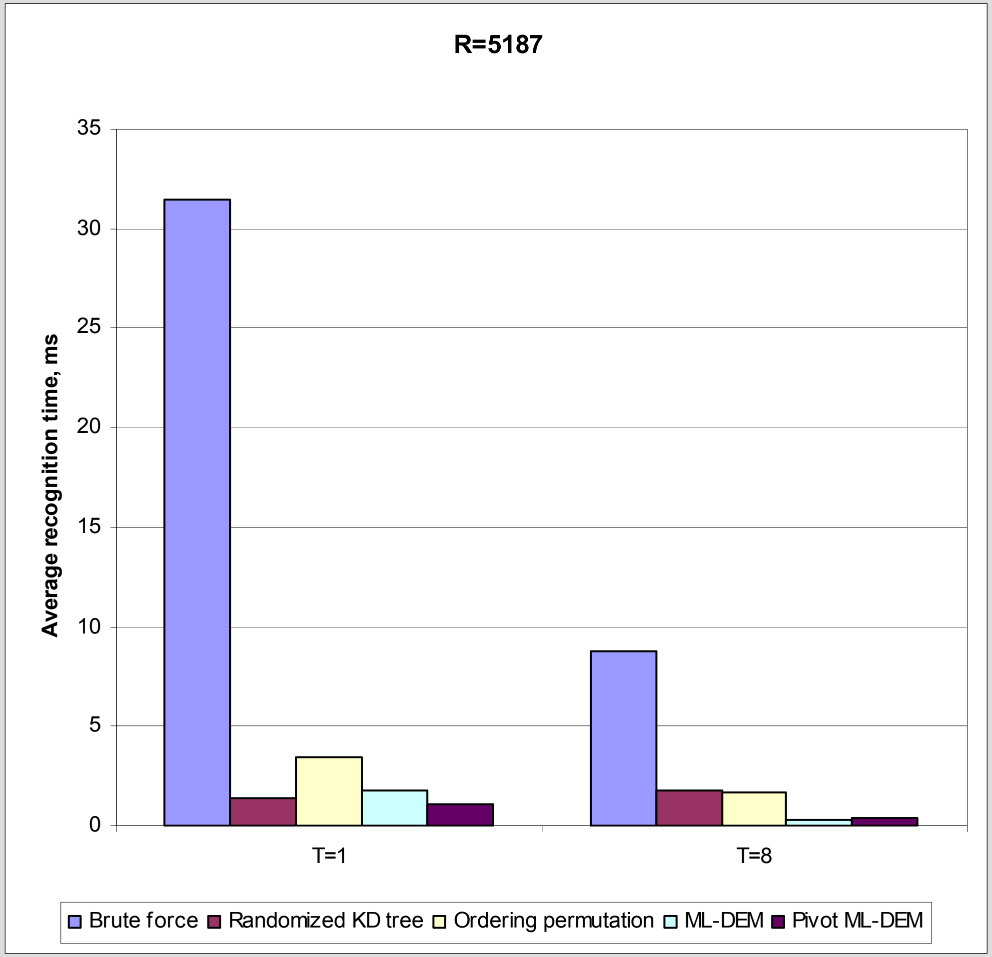
Experimental results (2). Essex dataset

Essex dataset (5187 photos of 323 persons in the training set, 1224 photos in the test set)

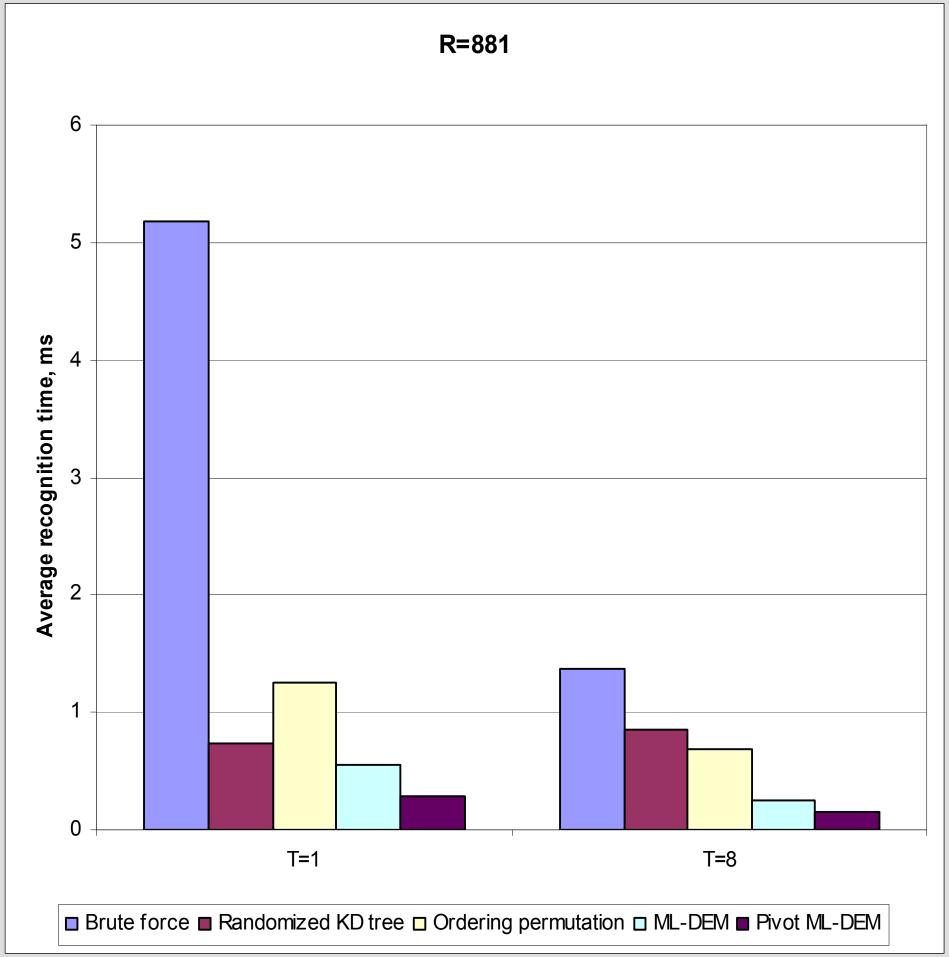
Clustering of the training set – 881 photos (medoids)

Average recognition time (ms.)

Original training set



Reduced training set



Error rate: 0.164%

0.573%

Simple example. DEM for students database

Distance matrix

	baranchikov	chuprakov	ershov	hohlova	kolchina	shal
baranchikov	0,000	0,610	0,473	0,686	0,643	0,608
chuprakov	0,536	0,000	0,481	0,561	0,639	0,611
ershov	0,505	0,593	0,000	0,583	0,580	0,514
hohlova	0,652	0,620	0,562	0,000	0,482	0,523
kolchina	0,656	0,655	0,557	0,486	0,000	0,483
shal	0,635	0,666	0,492	0,508	0,451	0,000

Threshold: 0,45

Distances to test image (Chuprakov)

baranchikov	chuprakov	ershov	hohlova	kolchina	shal
0,594	0,140	0,591	0,604	0,633	0,665

First model to check (randomly) “hohlova” (№ 4)

Model	$\varphi_\mu(4)$
baranchikov	0,010
chuprakov	0,003
ershov	0,001
hohlova	-
kolchina	0,029
shal	0,018

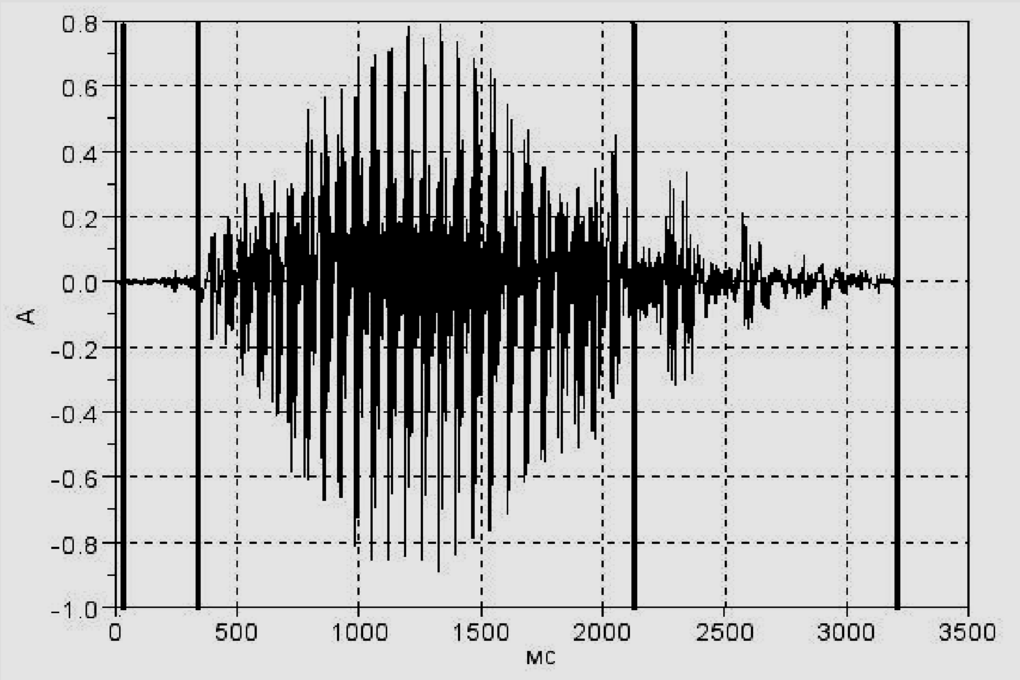
Second model to check “ershov” (№ 3)

Model	$\varphi_\mu(4)+\varphi_\mu(3)$
baranchikov	0,040
chuprakov	0,029
ershov	-
hohlova	-
kolchina	0,031
shal	0,038

Speech recognition. Simple features

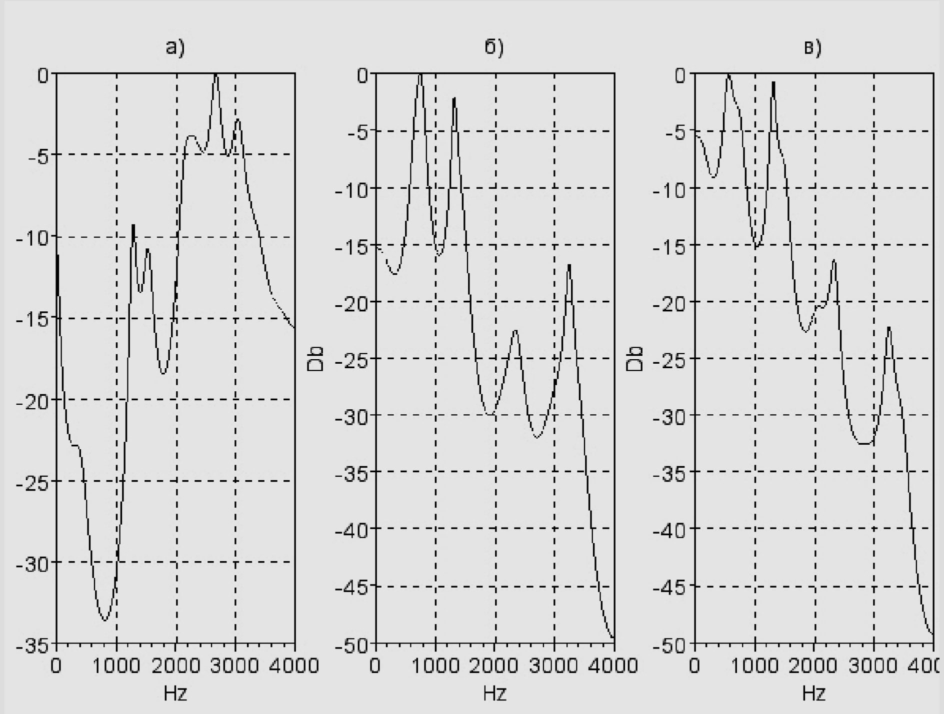
Word “шар” (ball, “sh aa r”)

Segmented speech signal



Sh aa r

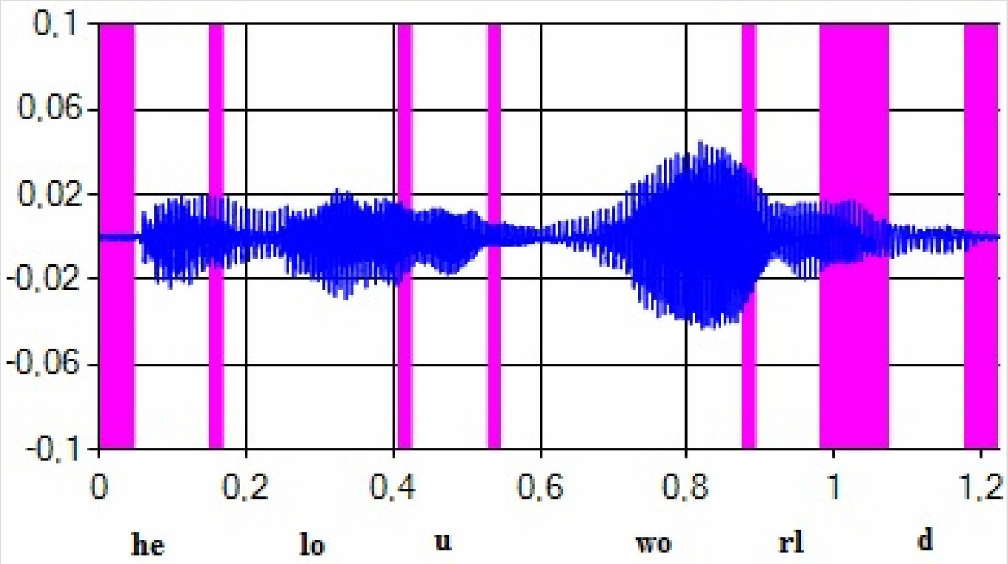
Power Spectral Densities (PSD) of each phone



Sh aa r

PSDs are usually estimated with the LPC (linear prediction coding) coefficients – autoregression model of the phone. Most popular is Levinson-Durbin procedure with the Burg (maximum entropy) method (see Marple, 1987)

Experimental results (3). Voice command recognition (a)



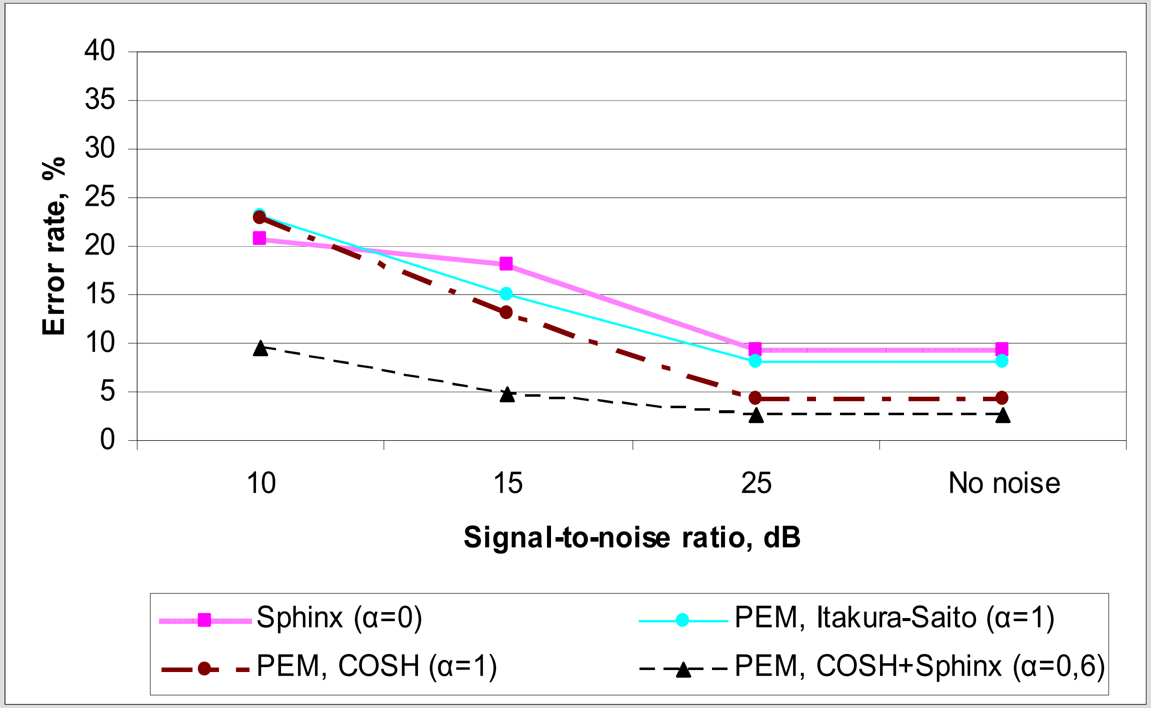
HT-PNN is implemented in voice control software in the **Phonetic Encoding Method (PEM)**. It requires the isolated syllable mode. We proposed to aggregate with fixed weight α the output of the PEM with the posterior probability of extracted syllable estimated by any ASR library (e.g. CMU Pocketsphinx). To train the system each speaker pronounce 10 vowels.

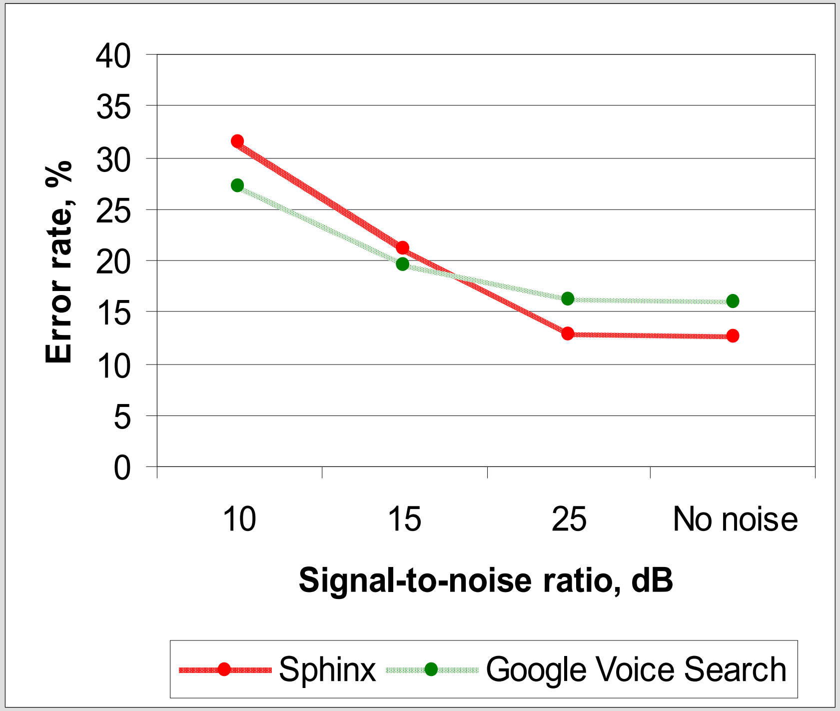
PNN	$\rho_{IS}(X_r(k_1), X(k)) = \frac{2}{F}\sum_{f=1}^{F/2}\left(\frac{G_{X(k)}(f)}{G_{X_r(k_1)}(f)} - \ln\frac{G_{X(k)}(f)}{G_{X_r(k_1)}(f)} - 1\right)$	Itakura-Saito divergence
HT-PNN	$\rho_{COSH}(X_r(k_1), X(k)) = \frac{1}{F}\sum_{f=1}^{F/2}\frac{\left(G_{X(k)}(f) - G_{X_r(k_1)}(f)\right)^2}{G_{X(k)}(f)\cdot G_{X_r(k_1)}(f)}$	COSH distance

Experimental results (3). Voice command recognition (b)

Parameters: frame length 30 ms, frame overlap 10 ms, autoregression model order p=12. Artificially generated white noise is added.

Error rates (%) for the list of 1830 Russian cities





1. Isolated syllable mode allow to not only extract voice commands but to increase the ASR accuracy
2. Addition of noise leads to high error rate of the original PEM without classifier fusion
3. Fusion of general ASR with the user vowel recognition is the best choice

Conclusion

We analyzed the methods of piecewise-regular object classification:

1. The dependence of the classifier choice on the number of classes and models in the database is highlighted.
2. Our brief survey showed that the current trends in the development of composite object recognition methods are connected with the refusal of complex algorithms of uncorrelated feature extraction and complication of the classifiers.
3. We emphasized one of the most exciting challenges in this field, namely, small number of models per each class. Most of recognition procedures in this case implement the nearest neighbor rule with various distances.
4. If the number of classes C is large, brute force solution is not computing efficient. Hence, approximate nearest neighbor algorithms can be applied.
5. Unfortunately, most of such algorithms allows to make the recognition faster only for very-large databases and do not work better than an exhaustive search for middle-sized databases. However, our experimental study show that there is the DEM which improves the recognition speed in several times.

Literature

Pattern recognition

1. S. Theodoridis, K. Koutroumbas, *Pattern Recognition*, 4th ed. Burlington, MA; London: Academic Press, 2008.
2. S. O. Haykin, *Neural Networks and Learning Machines*. 3th ed.. Harlow: Prentice Hall, 2008.
3. L. Rutkowski, *Computational Intelligence: Methods and Techniques*. Springer, 2010.

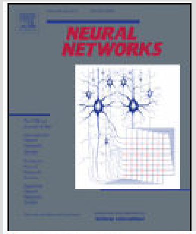
Image recognition

1. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", *in 2005 Proc. IEEE CVPR Conf.*, pp. 886–893.
2. D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, vol. 60, no. 2. pp. 91–110, 2004
3. P. Viola and M. Jones “Rapid object detection using a boosted cascade of simple features”, *in 2005 Proc. IEEE CVPR Conf.*, Vol. 1, pp. 511–518.
4. D. Cireşan, U. Meier, J. Masci and J. Schmidhuber, "Multi-column deep neural network for traffic sign classification", *Neural Networks*, vol. 32. pp. 333–338, 2012.

Speech recognition

1. J. Benesty, M. M. Sondhi and Y. Huang (eds.) *Springer Handbook of Speech Processing*. Berlin u.a.: Springer, 2008.
2. G. Hinton et al. "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups", *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.

See also



1. Savchenko A.V., Probabilistic neural network with homogeneity testing in recognition of discrete patterns set // Neural Networks. – 2013.– Vol. 46 (October).– pp. 227-241
2. Savchenko A.V., Directed enumeration method in image recognition // Pattern Recognition. – 2012. – 45(8) . – pp. 2952-2961.
3. Savchenko A.V., Khokhlova Ya.I. About Neural-Network Algorithms Application in Viseme Classification Problem With Face Video in Audiovisual Speech Recognition Systems, Optical Memory and Neural Networks (Information Optics), 2014, vol. 23, no. 1, pp. 34–42
4. Savchenko A.V. Phonetic encoding method in the isolated words recognition problem // Journal of Communications Technology and Electronics, 2014, Vol. 59, No. 4, pp. 310-315
5. Savchenko, A.V., Pattern recognition and increasing of the computational efficiency of a parallel realization of the probabilistic neural network with homogeneity testing, Optical Memory and Neural Networks (Information Optics), 2013, vol. 22, no. 3, pp. 184–192
6. Savchenko A.V. Phonetic words decoding software in the problem of Russian speech recognition // Automation and Remote Control, 2013, Vol. 74, No. 7, pp. 1225-1232
7. Savchenko, A.V., Adaptive Video Image Recognition System Using a Committee Machine, Optical Memory and Neural Networks (Information Optics), 2012, vol. 21, no. 4, pp. 219–226
8. Savchenko A.V. Real-Time Image Recognition with the Parallel Directed Enumeration Method // Proceedings of the Int.Conf on Vision Systems, M. Chen, B. Leibe, and B. Neumann (Eds.): ICVS 2013, LNCS 7963. St. Petersburg (Russia). 2013. pp. 123-132
9. Savchenko A.V. Nonlinear Transformation of the Distance Function in the Nearest Neighbor Image Recognition // In Proc. of International Conference on Computational Modeling of Objects Presented in Images, Y. Zhang and J.M.R.S. Tavares (Eds.): CompIMAGE 2014, LNCS 8641. Pittsburg (USA). 2014. pp. 261-266
10. Savchenko, A.V. Classification of a Sequence of Objects with the Fuzzy Decoding Method / A.V. Savchenko, L.V. Savchenko // In Proc. of International Conference on Rough Sets and Current Trends in Computing (RSCTC 2014). Granada & Madrid (Spain). C. Cornelis et al. (eds.): LNCS/LNAI.– 2014.– Vol. 8536.– P.309-318
11. Savchenko A.V. Semi-automated Speaker Adaptation: How to Control the Quality of Adaptation? // Proceedings of the Int.Conf on Image and Signal Processing, A. Elmoataz et al. (Eds.): ICISP 2014, LNCS 8509. Cherbourg (France). 2014. pp. 638-646
12. Savchenko A.V. Statistical Recognition of a Set of Patterns Using Novel Probability Neural Network // Proceedings of the Int.Conf on Artificial Neural Networks and Pattern Recognition: ANNPR-2012, LNCS/LNAI 7477. Trento (Italy). 2012. pp. 93-103
13. Savchenko A.V. Face Recognition in Real-Time Applications: Comparison of Directed Enumeration Method and K-d Trees // Proceedings of the Int.Conf on Perspectives in Business Informatics Research: BIR-2012, LNBIP 128. N.Novgorod (Russia). 2012. pp. 187-199

Thank you for listening!
Questions?