

A Study on Various Shape Descriptors and Matching Methods Applied in the General Shape Analysis

Abstract. In the paper, the General Shape Analysis problem is investigated using various combinations of shape descriptors and matching methods. Five shape descriptors were used, namely the Two-Dimensional Fourier Descriptor, Generic Fourier Descriptor, UNL-Fourier, Zernike Moments and Point Distance Histogram, and four matching methods - the Euclidean distance, Mahalanobis distance, correlation coefficient and C1 metric. The experiments made it possible to determine how matching methods influence the final effectiveness when a particular shape descriptor was applied.

Streszczenie. W artykule przedstawiono badanie różnorodnych kombinacji deskryptorów kształtu i metod dopasowania w problemie Ogólnej Analizy Kształtu. Wykorzystano pięć deskryptorów, a mianowicie Dwuwymiarowy Deskryptor Fouriera, Generic Fourier Descriptor, UNL-Fourier, Momenty Zernike'a oraz Point Distance Histogram, a także cztery metody dopasowania – odległość Euklidesową, odległość Mahalanobisa, współczynnik korelacji oraz metrykę C1. Eksperymenty pozwoliły zdeterminować, jaki wpływ mają metody dopasowania na ostateczną skuteczność eksperymentu. (Badanie różnych kombinacji deskryptorów kształtu i metod dopasowania w problemie Ogólnej Analizy Kształtu)

Keywords: General Shape Analysis, shape descriptors, matching methods

Słowa kluczowe: Ogólna Analiza Kształtu, deskryptory kształtu, metody dopasowania

Introduction

The General Shape Analysis (GSA) aims at finding one or a few most similar general templates for each test object, where a template is a simple geometrical figure, e.g. a triangle, rectangle or circle, and a test object is a more diversified shape. This approach enables to determine the most general and predominant shape features. The idea of the GSA is to represent all shapes using a particular shape description algorithm and calculate a similarity or dissimilarity between test objects and templates. Subsequently, a set of most similar templates indicated by the algorithm is compared with the results provided by people through inquiry forms concerning the same GSA task – the percentage convergence between the two gives the final effectiveness value of the experiment.

The General Shape Analysis was introduced in [3] and firstly applied for the Two-Dimensional Fourier Descriptor. In subsequent years, this approach has been examined with the use of other shape descriptors, among which were the UNL-Fourier descriptor [4], Generic Fourier Descriptor [5], Point Distance Histogram [4,5], Zernike Moments [6], Moment Invariants [6] or Curvature Scale Space [7]. According to the literature listed above, only the Euclidean distance was used as a matching method. The first application of the other shape matching method was presented in [8], where the correlation coefficient was applied to match Fourier-based shape representations. The GSA has been successfully applied in the identification of stamp types, which is useful in searching for presumably falsified digital documents [2]. The approach may also be applied in searching large multimedia databases where voice commands are used for shape retrieval [4].

The studies presented in the paper concern the investigation of various combinations of shape descriptors and matching methods. The shape descriptors include the Two-Dimensional Fourier Descriptor, Generic Fourier Descriptor, UNL-Fourier descriptor, Zernike Moments and Point Distance Histogram, and each of them is used to produce feature vectors of various size. For shape matching two dissimilarity measures were selected, namely the Euclidean distance and Mahalanobis distance, and two similarity measures – the correlation coefficient and C1 metric. An approach for estimating experimental effectiveness is as follows – if the first template indicated by the algorithm matches one out of three indications from the human benchmark result then the indicated template is considered proper. It needs to be emphasized that the GSA

is not concerned with studying the way in which people establish the similarity between some shapes, but it tries to find an appropriate substitute in the area of computer pattern recognition. Additionally, we should also think of how people describe things, because relatively often a shape of an unknown object is described using general and known features. Moreover, there is another approach to describe general shape features, introduced by Paul Rosin. The author investigated global shape measures to represent shape properties such as rectangularity, triangularity or ellipticity in a form of a single value [13,14].

The rest of the paper is organised as follows. The second section describes selected matching methods, i.e. methods for estimating similarity and dissimilarity between feature vectors. The third section briefly presents shape description algorithms selected for shape representation. The fourth section provides the description of the experiments and experimental results concerning the application of various combinations of shape descriptor and matching method in the GSA task. The last section summarizes and concludes the paper.

Shape Matching Methods

In the GSA, test objects are compared with the templates in order to estimate the similarity (or dissimilarity) between shapes. The similarity of shapes can be determined using measures based on the maximization of correlation between shapes, here the correlation coefficient and C1 metric are applied. In turn the dissimilarity measure is based on the minimization of the distance between shapes, and in the paper the Euclidean and Mahalanobis distances are used. All four selected matching methods are outlined below.

Let us take as an example two vectors $V_A(a_1, a_2, \dots, a_N)$ and $V_B(b_1, b_2, \dots, b_N)$ which represent object A and object B in a N -dimensional feature space. The Euclidean distance d_E between these two vectors is defined by means of the following formula [9]:

$$(1) \quad d_E(V_A, V_B) = \sqrt{\sum_{i=1}^N (a_i - b_i)^2}.$$

The Mahalanobis distance d_M between vectors V_A and V_B can be derived as follows [16]:

$$(2) \quad d_M(V_A, V_B) = \sqrt{(V_A - V_B)^T E^{-1} (V_A - V_B)},$$

where E^{-1} is the covariance matrix.

The correlation coefficient may be calculated both for the matrix and vector representations of a shape. The correlation between two matrices can be derived using the formula [1]:

$$(3) c_c = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \bar{A})^2\right) \left(\sum_m \sum_n (B_{mn} - \bar{B})^2\right)}}$$

where: A_{mn} , B_{mn} – pixel value with coordinates (m, n) , respectively in image A and B ; \bar{A} , \bar{B} – average value of all pixels, respectively in image A and B .

The C1 metric is also a similarity measure based on shape correlation. It is obtained by means of the following formula [11]:

$$(4) c_1(A, B) = 1 - \frac{\sum_{i=1}^H \sum_{j=1}^W |x_{ij} - y_{ij}|}{\sum_{i=1}^H \sum_{j=1}^W (|x_{ij}| + |y_{ij}|)}$$

where: A , B – matched shape representations; H , W – height and width of the representation.

Selected Shape Descriptors

The use of Fourier-based shape descriptors is widespread in pattern recognition thanks to its properties which include shape generalisation, robustness to noise, scale invariance and translation invariance. The Two-Dimensional Fourier Descriptor (2DFD) has the form of a matrix with absolute complex values, and is derived as follows [10]:

$$(5) C(k, l) = \frac{1}{HW} \left| \sum_{h=1}^H \sum_{w=1}^W P(h, w) \cdot e^{-i \frac{2\pi}{H} (k-1)(h-1)} \cdot e^{-i \frac{2\pi}{W} (l-1)(w-1)} \right|$$

where: H , W – height and width of the image in pixels; k – sampling rate in vertical direction ($k \geq 1$ and $k \leq H$); l – sampling rate in horizontal direction ($l \geq 1$ and $l \leq W$); $C(k, l)$ – value of the coefficient of discrete Fourier transform in the coefficient matrix in k row and l column; $P(h, w)$ – value in the image plane with coordinates h, w .

The UNL-Fourier (UNL-F) descriptor is composed of the UNL (named after Universidade Nova de Lisboa) descriptor and Fourier transform. The UNL utilizes a complex representation of Cartesian coordinates for points and parametric curves in discrete manner [12]:

$$(6) z(t) = (x_1 + t(x_2 - x_1)) + j(y_1 + t(y_2 - y_1)), \quad t \in (0, 1),$$

where $z_1 = x_1 + jy_1$ and $z_2 = x_2 + jy_2$ are complex numbers. In the next step, the centroid O is calculated [12]:

$$(7) O = (O_x, O_y) = \left(\frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n y_i \right),$$

and the maximal Euclidean distance between contour points and centroid is found [12]:

$$(8) M = \max_i \{ \|z_i(t) - O\| \} \quad \forall i = 1 \dots n \quad t \in (0, 1),$$

Based on the above formulations, a discrete version of the new coordinates is calculated as follows [12]:

$$(9) U(z(t)) = \frac{\left\| (x_1 + t(x_2 - x_1) - G_x) + j(y_1 + t(y_2 - y_1) - G_y) \right\|}{M} + j \times \operatorname{atan} \left(\frac{y_1 + t(y_2 - y_1) - G_y}{x_1 + t(x_2 - x_1) - G_x} \right).$$

Original pixels values are put into a square Cartesian matrix based on the new coordinates. This results in an image containing unfolded shape contour in polar coordinates, in which rows represent distances from the centroid and columns represent the angles. As a result, the 2DFD can be applied.

The Generic Fourier Descriptor (GFD) utilizes the transformation of a region shape to the polar coordinate system. It means that all pixel coordinates of an original image are transformed into polar coordinates. Next, the original pixel values are put to new coordinates on a rectangular Cartesian image, in which the row elements correspond to distances from the centroid and the columns to angles [15]. Again, the result is two-dimensional and the Fourier transform can be applied.

The Point Distance Histogram (PDH) is a shape descriptor that utilizes information about the shape contour. In order to derive a PDH representation, an origin of the polar transform of a contour is firstly selected, usually a centroid O . Polar coordinates are stored in two vectors – θ^i for angles and P^i for radii [5]:

$$(10) \theta_i = a \tan \left(\frac{y_i - O_y}{x_i - O_x} \right), \rho_i = \sqrt{(x_i - O_x)^2 + (y_i - O_y)^2}.$$

In the next step, the values in θ^i are converted to the nearest integers. The elements in θ^i and P^i are rearranged with respect to the increasing values in θ^i and denoted as θ^j , P^j . If there are any equal elements in θ^j , then only the element with the highest value P^j is left. This results in a vector which has no more than 360 elements. Next, only the P^j vector is selected for further processing and denoted as P^k , where $k=1, 2, \dots, m$ and $m \leq 360$. The elements of P^k vector are normalized to ρ_k and assigned to bins in the histogram (ρ_k to l_k) [5]:

$$(11) l_k = \begin{cases} r, & \text{if } \rho_k = 1 \\ \lfloor r \rho_k \rfloor, & \text{if } \rho_k \neq 1 \end{cases}$$

where r is a previously determined number of bins. In the next step, the values in the histogram bins are normalized according to the highest one. Ultimately, the final histogram which represents a shape is obtained and can be written as the following function $h(l_k)$ [5]:

$$(12) h(l_k) = \sum_{k=1}^m b(k, l_k),$$

where [5]:

$$(13) b(k, l_k) = \begin{cases} 1, & \text{if } k = l_k \\ 0, & \text{if } k \neq l_k \end{cases}$$

Zernike Moments (ZM) are orthogonal moments. Among the advantages of this descriptor are rotation invariance, and robustness to noise and minor variations in shape. The complex Zernike Moments are derived from orthogonal Zernike polynomials, which are a set of functions orthogonal over the unit disk $x^2 + y^2 = 1$. The Zernike Moments of order n and repetition m of a region shape $f(x, y)$ can be obtained by the following formula [15]:

$$(14) Z_{nm} = \frac{n+1}{\pi} \sum_r \sum_{\theta} f(r \cos \theta, r \sin \theta) \cdot R_{nm}(r) \cdot \exp(jm\theta),$$

where $r \leq 1$ and $R_{nm}(r)$ is the orthogonal radial polynomial [15]:

$$(15) R_{nm}(r) = \sum_{s=0}^{(n-m)/2} (-1)^s \frac{(n-s)!}{s! \times \left(\frac{n-2s+|m|}{2} \right)! \left(\frac{n-2s-|m|}{2} \right)!} r^{n-2s},$$

where $n=0,1,2,\dots; 0 \leq m \leq n; n-|m|$ is even.

The Description of the Experiments and Experimental Results

During the experiments, five different shape descriptors and four matching methods were used. The database used in the experiments is depicted in Fig. 1 and consisted of 200x200 pixel size images with white backgrounds and black silhouettes. Each experiment investigated one combination of a shape descriptor and matching method. In the first step, all shapes were represented using a selected variant of the shape descriptor that varied significantly in terms of size. In case of shape descriptors based on the Fourier transform, various parts of the original absolute spectrum were investigated, namely 2x2, 5x5, 10x10, 25x25 and 50x50 subparts of the coefficient matrix. Each block was transformed into a vector to form a final shape representation. The Zernike Moments descriptor was calculated for orders from 1 to 20, what resulted in feature vectors having from 2 to 121 elements. The Point Distance Histogram descriptor had seven variants that were obtained for 2, 5, 10, 25, 50, 75 and 100 histogram bins, and simultaneously produced feature vectors of size equal to the number of bins.

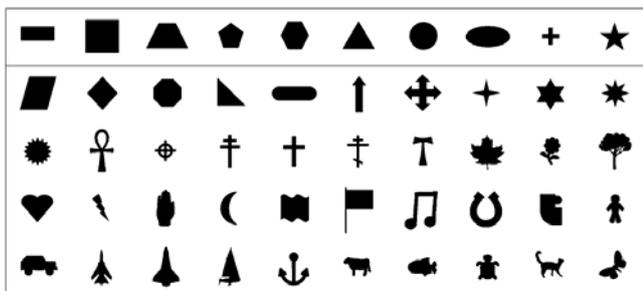


Fig. 1 Shapes used in the experiments divided into 10 templates (first row) and 40 test objects (rest) [5].

In the second step, the representations of test objects were matched with the representations of templates by calculating the similarity or dissimilarity measure. Lastly, one most similar template was selected for each investigated object, giving a set of templates. The effectiveness of the experiment was estimated by calculating the percentage of the templates selected in the experiment that was consistent with the templates indicated by people in the inquiries concerning the same GSA task. The main goal of the experiments was to select the combination of a shape descriptor and matching method that gives the highest effectiveness and, additionally, in the case of several combinations with the same percentage effectiveness, in which the size of the shape representation is the smallest. The following part of this section describes the experimental results.

The first set of the experiments utilized the Two-Dimensional Fourier Descriptor and five different absolute spectrum subparts. The percentage effectiveness values for each combination of a shape descriptor and matching method are provided in Fig. 2. As can be seen in Fig. 2, the effectiveness values vary significantly and the weakest results were achieved in case of the use of the Mahalanobis distance. The highest effectiveness was obtained in the case of combinations with the percentage value equal to 55%. The best result can be attributed to the 5x5 subpart of the 2DFD and both similarity measures – correlation coefficient and C1 metric.

In the second set of the experiments, the Generic Fourier Descriptor was used and again five absolute spectrum subparts were investigated (see Fig. 3).

Compared to the previous experiment, the best result was obtained using a dissimilarity measure – the Euclidean distance, and the smallest feature vector – 2x2 subpart of the absolute spectrum. Similarly as in the previous case, the Mahalanobis distance provided the lowest effectiveness values.

Two-Dimensional Fourier Descriptor

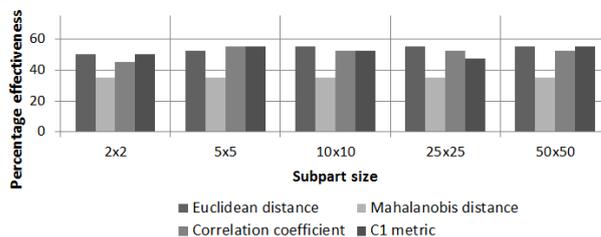


Fig. 2 Bar chart representing the experimental results obtained using the 2DFD.

Generic Fourier Descriptor

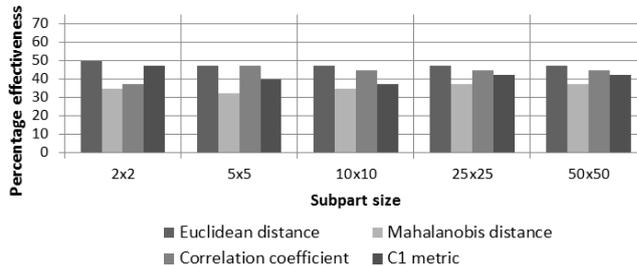


Fig. 3 Bar chart representing the experimental results obtained using the GFD.

UNL-Fourier Descriptor

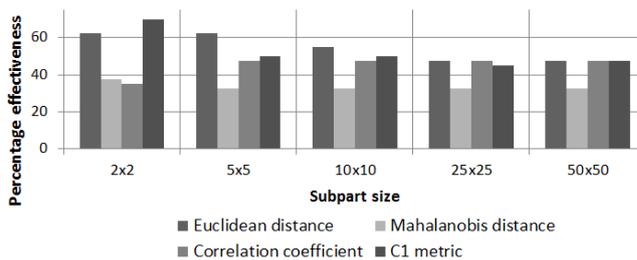


Fig. 4 Bar chart representing the results obtained using the UNL-F.

The third set of the experiments included the application of the UNL-Fourier descriptor and again various subparts of the Fourier coefficient matrix. The results are provided in Fig. 4. Three combinations stood out – 2x2 and 5x5 subparts of the UNL-F, which were matched using Euclidean distance, and 2x2 subpart of the UNL-F matched using C1 metric. These combinations gave 62,5% twice and 70% respectively. It is worth noting that the smallest feature vectors were sufficient to indicate the templates consistent with those selected by people in the inquiries.

Zernike moments

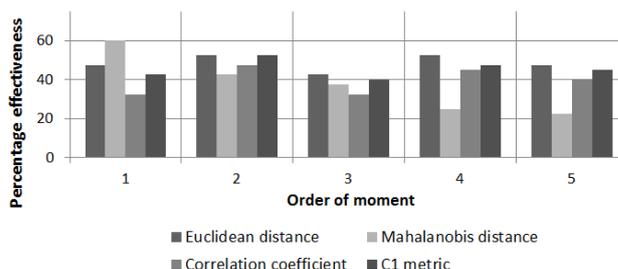


Fig. 5 Bar chart representing the experimental results obtained using the ZM.

The fourth set of experiments concerned the investigation of Zernike Moments descriptor and different orders of moment were used. The results are varied – the percentage effectiveness values range from 22.5% to 60%. Surprisingly, the best results were observed when the Mahalanobis distance was applied as the matching method and the first-order moment was used. In this case the feature vector had only two elements. Fig. 5 presents selected results obtained using ZM.

The last set of the experiments examined the Point Distance Histogram descriptor. A different number of histogram bins was utilized, what resulted in a varying number of elements in each feature vector. As can be seen in Fig. 6, the highest effectiveness value was equal to 50% and was obtained for the combination of the PDH descriptor calculated for five histogram bins and C1 metric.

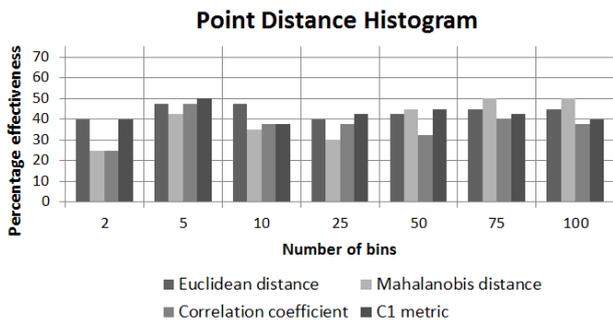


Fig. 6 Bar chart representing the experimental results obtained using the PDH.

Summary and Conclusions

In the paper, some exemplary solutions to the problem of the General Shape Analysis were investigated. In solving the GSA problem we are establishing the degree of similarity between test objects and general templates – one or few most similar templates are selected and compared with benchmark results in order to estimate the effectiveness of the experiment. The main goal of the experiments was to examine various combinations of shape descriptors and matching methods. Five shape descriptors were used to calculate shape representations (feature vectors) of various size. The descriptors comprised the Two-Dimensional Fourier Descriptor, Generic Fourier Descriptor, UNL-Fourier, Zernike Moments and Point Distance Histogram. The matching methods included two similarity measures, namely the correlation coefficient and C1 metric, and two dissimilarity measures – the Euclidean and Mahalanobis distances.

Based on the experimental results, the best solution for the GSA problem was selected, i.e. a combination of a shape descriptor and matching method which gave the highest percentage effectiveness and when the smallest feature vector was used. Accordingly, the best solution for the GSA problem is the combination of the UNL-Fourier descriptor, 2x2 subpart of the absolute spectrum and C1 metric. Additionally, both the calculation of description vectors and similarity measures between shapes are not time-consuming. There are slight differences between runtimes when using various matching methods and previously calculated descriptors, however they are not significant for small-sized description vectors.

By way of conclusion, it needs to be highlighted that the matching method has a significant impact on the final effectiveness of the experiment. The effectiveness values also depend on the applied version of the shape descriptor. Therefore, taking into consideration one particular shape description algorithm, each combination of a feature vector and matching method produces different experimental

results. This in turn may indicate that some feature vectors represent more significant shape features, enabling easy recognition and matching of all shapes with common general characteristics. It should be emphasized that matching method does not change the original efficiency of the shape description algorithm. A high diversity in effectiveness values stems from the fact that each matching method is based on different inputs, therefore it should be properly selected to fit the actual problem and the shape descriptor applied. Summarizing, three factors can affect the final experimental result: a shape description algorithm, the size of a feature vector and a method for estimating similarity or dissimilarity between shape representations.

REFERENCES

- [1] Chwastek T., Mikrut S., The problem of automatic measurement of fiducial mark on air images (in polish), *Archives of Photogrammetry, Cartography and Remote Sensing*, 16 (2006) 125-133
- [2] Forczmański P., Frejlichowski D., Robust stamps detection and classification by means of general shape analysis. In: Bolc L. et al (Eds.) *ICCVG 2010. Lect. Notes Comput. Sc.*, 6374 (2010), 360-367
- [3] Frejlichowski D., General shape analysis using fourier shape descriptors. In: Świątek J. et al. (Eds.) *Information Systems Architecture and Technology – System Analysis in Decision Aided Problems* (2009), 143-154
- [4] Frejlichowski D., An Experimental Comparison of Seven Shape Descriptors in the General Shape Analysis Problem. In: Campilho A., Kamel M. (Eds.) *ICIAR 2010, Part I. Lect. Notes Comput. Sc.*, 6111 (2010), 294-305
- [5] Frejlichowski D., An experimental comparison of three polar shape descriptors in the general shape analysis problem. In: Świątek J. et al. (Eds.) *Information Systems Architecture and Technology – System Analysis in Decision Aided Problems* (2010), 139-150
- [6] Frejlichowski D., The Application of the Zernike Moments to the Problem of General Shape Analysis, *Control Cybern.*, 40 (2011), No. 2, 515-526
- [7] Frejlichowski D., Application of the curvature scale space descriptor to the problem of general shape analysis, *Przełąd Elektrotechniczny*, 10b/2012 (2012), 209-212
- [8] Frejlichowski D., Gościewska K., Application of 2D Fourier Descriptors and Similarity Measures to the General Shape Analysis Problem, In: Bolc L. et al. (Eds.) *ICCVG 2012. Lect. Notes Comput. Sc.*, 7594 (2012), 371-378
- [9] Kpalma K., Ronsin J., An overview of advances of pattern recognition systems in computer vision. In: Obinata G., Dutta A. (Eds.) *Vision Systems: Segmentation and Pattern Recognition* (2007), 169-194
- [10] Kukharev G., *Digital Image Processing and Analysis*. SUT Press, Stettin, Poland (1998)
- [11] Lam K.M., Yan H., An analytic-to-holistic approach for face recognition based on a single frontal view, *IEEE T. Pattern Anal.* 20 (1998), no. 7, 673-686
- [12] Rauber T. W., Two Dimensional shape description. Technical report: GR UNINOVA-RT-10-94. Universidade Nova de Lisboa, Lisboa, Portugal (1994)
- [13] Rosin P. L.: Measuring shape: ellipticity, rectangularity, and triangularity, *Mach. Vision Appl.*, 14 (2003), 172-184
- [14] Rosin P. L., Computing global shape measures. In: Chen C.H. et al. (Eds.) *Handbook of Pattern Recognition and Computer Vision*, 3rd ed. (2005), 177-196
- [15] Yang M., Kpalma K., Ronsin J., A survey of shape feature extraction techniques. In: Yin P.Y. (Ed.) *Pattern Recogn.* (2008), 43-90
- [16] Zhang D., Image retrieval based on shape. Dissertation, Faculty of Information Technology, Monash University, Australia (2002)

Authors: Dariusz Frejlichowski, Ph.D., D.Sc., Eng., E-mail: dfrejlichowski@wi.zut.edu.pl; Katarzyna Gościewska, M.Sc., Eng., E-mail: kgosciewska@wi.zut.edu.pl; West Pomeranian University of Technology, Szczecin, Faculty of Computer Science and Information Technology, Żołnierska 52, 71-210, Szczecin, Poland.