IDIAP RESEARCH REPORT

Local Binary Patterns as an Image Preprocessing for Face Authentication

Guillaume Heusch [a] Yann Rodriguez [a]
Sébastien Marcel [a]

IDIAP–RR 05-76

November 2005

[a] IDIAP Research Institute

IDIAP Research Institute **www.idiap.ch**
Rue du Simplon 4 P.O. Box 592 1920 Martigny – Switzerland
Tel: +41 27 721 77 11 Fax: +41 27 721 77 12 Email: info@idiap.ch

IDIAP Research Report 05-76

Local Binary Patterns as an Image Preprocessing for Face Authentication

Guillaume Heusch   Yann Rodriguez   Sébastien Marcel

November 2005

Abstract. One of the major problem in face authentication systems is to deal with variations in illumination. In a realistic scenario, it is very likely that the lighting conditions of the probe image does not correspond to those of the gallery image, hence there is a need to handle such variations. In this work, we present a new preprocessing algorithm based on Local Binary Patterns (LBP): a texture representation is derived from the input face image before being forwarded to the classifier. The efficiency of the proposed approach is empirically demonstrated using both an appearance-based (LDA) and a feature-based (HMM) face authentication systems on two databases: BANCA and XM2VTS (with its darkened set). Conducted experiments show a significant improvement in terms of verification error rates and compare to results obtained with state-of-the-art preprocessing techniques.

1 Introduction

Since over 30 years, face recognition has been an active research area and different systems are now capable to properly recognize people under controlled environment (i.e. when the probe and the gallery images are acquired under the same global viewing conditions) [24]. Nevertheless, there are still open challenges such as recognition with varying pose, expression or illumination. In this paper, we will focus on the latter. Existing algorithms to cope with changes in lighting conditions can be divided into two main categories: model-based and preprocessing.

Model-based approaches derive a *model* of an individual face, which will account for variation in lighting conditions. Such algorithms relies on the assumption that the collection of face images under any illumination condition (the so-called face space) is a low-dimensional subspace of the image space. Several algorithms were developed based on this idea: examples include the spherical harmonics representation [2] and the illumination cone [8]. However, such algorithms require a training set with several different lighting conditions for the same subject, which can be considered as a weakness for realistic applications.

On the other hand, the preprocessing approach tries to minimize the effect of illumination variations prior to classification. Compared to the first approach, preprocessing has two main advantages: it is completely stand-alone and thus can be used with any classifier. Moreover, no training images are required to model illumination variations, since such algorithms turn the original image into another representation, which is supposed to be less sensitive to such variations. A recent study of preprocessing algorithms applied to face authentication [19] showed the superiority of the model proposed by Gross & Brajovic [9]. However, this approach involves the resolution of a partial differential equation, which is either computationaly expensive (brute force resolution) or approximate (iterative methods). Moreover, the algorithm depends on a normalization parameter and hence face authentication cannot be performed in a fully automated way, since this parameter has to be manually set *a priori*.

In this work, we propose to use Local Binary Patterns (LBP) [17] as a preprocessing step prior to the face authentication task. The LBP is a nonparametric and computationnaly simple descriptor of local texture patterns. Moreover, it is invariant to monotonic grayscale transformation, hence the LBP representation may be less sensitive to changes in illumination.

The remainder of this paper is organized as follows: in the next section, we briefly describe the normalization proposed in [9] and present our approach based on LBP. Section 3 introduces the task of face authentication and presents the systems we used in our experiments. Experimental setup as well as authentication results are presented in section 4 and a conclusion is drawn in section 5.

2 Illumination Normalization

The aim of this section is to describe the preprocessing steps we investigate in our experiments. As a baseline, the well-known histogram equalization was used prior to classification. The remainder of the section presents the state-of-the-art algorithm developed by Gross & Brajovic [9], as well as the approach we propose, based on Local Binary Patterns.

2.1 Gross & Brajovic Model

In their paper [9], Gross and Brajovic addressed the problem of lighting normalization by using calculus of variations. Recovering the reflectance from the illuminance-reflectance model, where an image is regarded as the product between the illuminance and the reflectance: $I(x, y) = R(x, y)I(x, y)$, is an ill-posed problem. Based on the common assumption that the luminance $L(x, y)$ can be estimated as a blurred version of the original image, they derive the following energy-based model to recover $L(x, y)$:

$$E(L) = \iint \rho(x, y)(L - I))^2 dxdy + \lambda \iint (L_x^2 + L_y^2) dxdy \quad (1)$$

where the first term forces the luminance function to be close to the image and the second term adds a smoothness constraint on $L(x, y)$. $\rho(x, y)$ are the anisotropic diffusion coefficients and the parameter λ controls the relative importance of the smoothness constraint. Note also that if $\rho(x, y) = 1 \, \forall \, x, y$, then isotropic diffusion is performed. The minimum of this functional is found using the corresponding discretized Euler-Lagrange equation:

$$\begin{aligned} L_{i,j} \quad &+\lambda \Big[\tfrac{1}{\rho_{i,j-}}(L_{i,j} - L_{i,j-1}) + \tfrac{1}{\rho_{i,j+}}(L_{i,j} - L_{i,j+1}) \\ &+\tfrac{1}{\rho_{i-,j}}(L_{i,j} - L_{i-1,j}) + \tfrac{1}{\rho_{i+,j}}(L_{i,j} - L_{i+1,j}) \Big] \quad = I_{i,j} \end{aligned} \tag{2}$$

where $\frac{1}{\rho_{a,b}}$ denotes the anisotropic coefficients and are defined as the inverse of Weber's contrast (equation 3). The resolution of equation (2) is carried out using iterative methods [11].

$$\rho_{a,b} = \frac{|I_a - I_b|}{\min(I_a, I_b)} \tag{3}$$

2.2 Our Approach: Local Binary Patterns

The local binary pattern (LBP) is a non-parametric operator which describes the local spatial structure of an image. Ojala et al. [17] first introduced this operator and showed its high discriminative power for texture classification. At a given pixel position (x_c, y_c), LBP is defined as an ordered set of binary comparisons of pixel intensities between the center pixel and its eight surrounding pixels (Figure 1). The decimal form of the resulting 8-bit word (LBP code) can be expressed as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c) 2^n \tag{4}$$

where i_c corresponds to the grey value of the center pixel (x_c, y_c), i_n to the grey values of the 8 surrounding pixels, and function $s(x)$ is defined as:

$$s(x) = \begin{cases} 1 & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \, . \end{cases} \tag{5}$$

By definition, the LBP operator is unaffected by any monotonic gray-scale transformation which

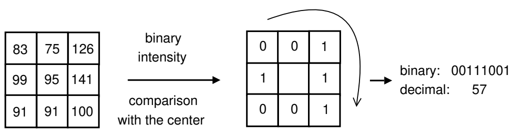


Figure 1: The LBP operator.

preserves the pixel intensity order in a local neighbourhood. Note that each bit of the LBP code has the same significance level and that two successive bit values may have a totally different meaning. Actually, The LBP code may be interpreted as a kernel structure index.

Later, Ojala et al. [18] extended their original LBP operator to a circular neighbourhood of different radius size. Their $LBP_{P,R}$ notation refers to P equally spaced pixels on a circle of radius R. In this paper, we use the $LBP_{8,2}$ operator which is illustrated in Figure 2.

Due to its texture discriminative property and its very low computational cost, LBP is becoming very popular in pattern recognition. Recently, LBP has been applied for instance to face detection [12],

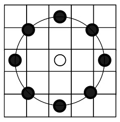


Figure 2: The extended LBP operator with (8,2) neighbourhood. Pixel values are interpolated for points which are not in the center of a pixel.

face recognition [23, 1], image retrieval [21], motion detection [10] or visual inspection [22]. In these previous works an image was usually divided in small regions. For each region, a cumulative histogram of LBP codes, computed at each pixel location within the region, was used as a feature vector. However, in the present work, we do not use LBP to perform feature extraction, but to preprocess the input image: the face is represented with its texture patterns given by the LBP operator at each pixel location (Figure 3).

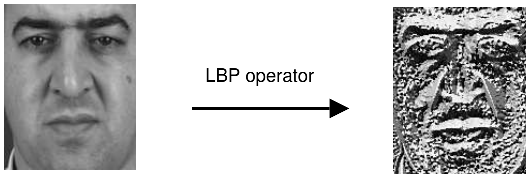


Figure 3: Original image (left) processed by the LBP operator (right).

3 Face Authentication

The task of face authentication is to make a decision, based on the face image of a client C and on a claimed identity: either the claimed identity corresponds to the identity derived from the face image (*true claimant*), or the client is an *impostor*. Hence, a face authentication system has to give an opinion on whether the client is the true claimant or not, given its face image. In this section, we present the two different face authentication algorithms used in our experiments. The first one is based on linear discriminant analysis (LDA) and is said to be holistic: the whole face is treated as a single vector. On the contrary, the second one is based on local features and Hidden Markov Models (HMM).

3.1 PCA-LDA

Linear discriminant analysis was first applied in face recognition in [3] and has been successfully applied to the problem of face authentication [13] [14]. The LDA algorithm is based on Fisher's linear discriminant formula to find the optimal (discriminative) subspace where data will be projected. Let us consider a training set composed of N examples x_i where each example is labeled and belongs to one of C classes. The training data are used to derive the between-class scatter matrix S_b and the within-class scatter matrix S_w:

$$S_b = \sum_{i=1}^{C} N_i\,(\mu_i - \mu)(\mu_i - \mu)^T \tag{6}$$

$$S_w = \sum_{i=1}^{C} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i^T) \tag{7}$$

where N_i are the number of examples in class i, μ_i is the mean of the examples in class i and μ is the mean of all training data. Finally, the optimal projection W_{lda} is the orthogonal matrix satisfying equation (8):

$$W_{lda} = \underset{W}{\operatorname{argmax}} \frac{|W^T S_b W|}{|W^T S_w W|} \tag{8}$$

In practice, Principal component analysis (PCA) is first applied on the data so as to achieve decorrelation and dimensionality reduction. The projected face representations into the coordinate system of eigenvectors are then used as training examples to derive the optimal projection W_{lda}.

Authentication Considering the holistic face representation as a vector denoted by x, it is linearly projected to obtain the feature vector y:

$$y = W_{lda}^T W_{pca}^T x$$

where W_{pca}^T is the PCA matrix, containing the eigenvectors of the training data covariance matrix. Finally, classification is performed using a metric in the feature space: considering two feature vectors, a template y_t and a sample y_s, their correlation is computed according to the following formula:

$$\operatorname{corr}(y_s, y_t) = 1 - \frac{< y_t, y_s >}{\|y_t\| \cdot \|y_s\|}$$

The decision on the identity claim is based on a threshold τ: the claim is accepted if $\operatorname{corr}(y_s, y_t) \leq \tau$ and rejected otherwise.

3.2 1D-HMM

It has been shown that face authentication systems based on statistical models are suitable for the face authentication task [6]. On the contrary to the LDA-based system where a face image is treated as a whole, such models generally uses local features to analyze a given face image. Several statistical models have been proposed to model the face representation, examples include Gaussian Mixtures Models (GMM) [6], Hidden Markov Models (HMM) [5] and Bayesian networks [16].

With the one dimensional HMM (1D-HMM) structure proposed in [4] the face representation is analyzed on a block-by-block basis: from each block, discrete cosine transform (DCT) coefficients are obtained; the first M coefficients compose the feature vector corresponding to the block under consideration. The main advantage of this topology is to relax the rigid horizontal constraints present in the traditional 1D-HMM, where an entire strip is used as observation.

Such a model is characterised by its number of states N_s, its state transition matrix A and its emission probability distributions B, which are considered to be a mixture of gaussians. Let us define X as the observation sequence, representing the consecutive horizontal strips of the face: $X = \{x_s\}_{s=1}^{N_s}$. Each strip is itself a sequence of N_B observation vectors $x_s = \{x_b\}_{b=1}^{N_B}$ representing the consecutive blocks composing the strip. Under the assumption that the blocks composing a strip are independent and identically distributed, the likelihood of the strip s for the state S_j is given by:

$$P(\vec{x}_s | q_s = S_j) = \prod_{b=1}^{N_B} P(\vec{x}_s^b | \lambda_j) \tag{9}$$

Then, the likelihood of an observation sequence X is:

$$P(X|\lambda) = \sum_{\forall Q} \prod_{s=1}^{N_S} \prod_{b=1}^{N_B} P(\vec{x}_s^b | \lambda_j) \prod_{s=2}^{N_S} a_{q_{s-1}, q_s} \tag{10}$$

where:

$$P(\vec{x}_b|\lambda_j) = \sum_{k=1}^{N_G} m_k^s \, \mathcal{N}(\vec{x}_s^b|\vec{\mu}_k^s, \Sigma_k^s) \quad (11)$$

The training phase is performed in two stages: first, a so-called *world-model* is learned using image data coming from various identities. Then, each client model is *adapted* from the world-model through maximum *a posteriori* (MAP) adaptation [5].

Authentication Let us denote $\lambda_{\bar{C}}$ as the parameter set defining the world-model whereas λ_C represents the client-specific parameters. Given a client claim and its face representation X, an opinion on the claim is given by the following likelihood ratio:

$$\Lambda(X) = \log P(X|\lambda_C) - \log P(X|\lambda_{\overline{C}}) \quad (12)$$

where $P(X|\lambda_C)$ is the likelihood of the claim coming from the true client and $P(X|\lambda_{\overline{C}})$ is an approximation of the likelihood of the claim coming from an impostor. Based on a threshold τ, the claim is accepted if $\Lambda(X) \geq \tau$ and rejected otherwise.

4 Experiments & Results

Face windows are manually extracted from the original images based on the eyes position. They are resized to 64x80 pixels, converted to grayscale and finally treated by the investigated preprocessing algorithms before training and testing. Tables 1 and 2 presents results for each face authentication systems with three preprocessing techniques: histogram equalization (denoted as HEQ), the normalization model proposed in [9] (denoted as GROSS) and our approach (denoted as LBP). Performance are reported in terms of *Half Total Error Rate* (HTER), which combines the *False Acceptance Rate* and the *False Rejection Rate*:

$$HTER = \frac{(FAR + FRR)}{2} \quad [\%] \quad (13)$$

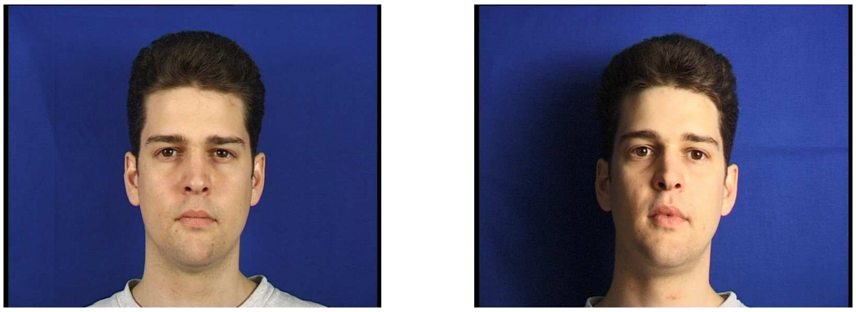
4.1 XM2VTS

The XM2VTS database [15] is a multi-modal database containing 295 idendities, among which 200 are used as true clients (the remainder are considered as impostors). Recordings were acquired during four sessions over a period of five months under controlled conditions (blue background, uniform illumination) for the standard set. The darkened set contains four images of each subject acquired with side lighting. Figure 4 shows images coming from both sets of the database.

Together with the standard set, there exists two detailed protocols used to perform authentication experiments [15]. Concerning darkened set experiments, protocols are the same but for the testing phase: it is done on the darkened images. Table 1 shows the results obtained on the XM2VTS database, where LP1 and LP2 refers to the Lausanne Protocol 1 and 2 respectively.

4.2 BANCA

The BANCA database [7] was especially meant for multi-modal biometric authentication and contains 52 clients (English corpus), equally divided into two groups $g1$ and $g2$ used for development and evaluation respectively. Each corpus is extended with an additional set of 30 other subjects and is referred as the *world model*. Image acquisition was performed with two different cameras: a cheap analogue webcam, and a high-quality digital camera, under several realistic scenarios: controlled (high-quality camera, uniform background, controlled lighting), degraded (webcam, non-uniform background) and adverse (high-quality camera, arbitrary conditions). Figure 5 shows examples of the different acquisition scenarios.



(a) standard set (b) darkened set

Figure 4: examples of images in the XM2VTS database.

	standard		darkened	
FA system	LP1	LP2	LP1	LP2
LDA HEQ	**2.97**	**0.84**	10.86	17.02
LDA GROSS	5.76	4.88	12.62	13.38
LDA LBP	4.56	1.43	**9.110**	**10.44**
HMM HEQ	2.04	1.40	37.32	37.54
HMM GROSS	5.53	4.18	12.01	11.96
HMM LBP	**1.37**	**0.97**	**9.61**	**9.88**

Table 1: Performances on the standard and the darkened sets of the XM2VTS database.

(a) controlled (b) degraded (c) adverse

Figure 5: Example of the different scenarios in the BANCA database

In the BANCA protocol, seven distinct configurations for the training and testing policy have been defined. In our experiments, the configurations referred as Match Controlled (Mc), Unmatched Adverse (Ua), Unmatched Degraded (Ud) and Pooled Test (P) are used. All of the of listed configurations use the same training conditions: each client is trained using images from the first recording session, which corresponds to the controlled scenario. Testing is then performed on images taken from the controlled scenario (Mc), adverse scenario (Ua), degraded scenario (Ud), while (P) does the test for each of the previously described configurations. Table 2 shows the results obtained on the different protocols of the BANCA database.

FA system	Mc	Ua	Ud	P
LDA HEQ	**3.75**	20.13	**14.46**	15.52
LDA GROSS	3.97	**17.40**	15.01	**14.24**
LDA LBP	5.83	19.52	15.61	16.30
HMM HEQ	2.40	19.87	18.75	18.32
HMM GROSS	**1.92**	**11.70**	**7.21**	11.75
HMM LBP	2.40	15.06	9.93	**11.70**

Table 2: Performance on the different protocols of the BANCA database.

4.3 Discussion

Firstly, results on the XM2VTS database show that the LBP representation is suitable when there is a strong mismatch, in terms of illumination conditions, between the gallery and the probe image. This is evidenced by experiments on the darkened set, where the error rates of both classifiers are decreased when using the LBP representation. Moreover, this texture representation outperforms the illumination normalization approach (GROSS). Interestingly, standard experiments also show an improvement for the HMM-based classifier: this suggest that our preprocessing technique is well suited for feature-based approaches. Although the best results obtained with LDA are with the use of histogram equalization, error rates of the LBP are still lower than the GROSS normalization.

On the BANCA database, the LDA classifier seems to have a good discriminative capability, since none of the method clearly outperforms the others (although GROSS normalization is the best). A possible explanation could reside in the fact that we use the Spanish corpus (with all scenarios) to train the LDA, hence it may capture by itself the changes in acquisition conditions. Concerning the HMM-based classifier, GROSS normalization results are better for three of the four investigated protocol, and reduces error rates by a significant amount compared to histogram equalization. Results obtained with the LBP representation are comparable, although performances are a bit worse.

To summarize, conducted experiments shows that the proposed preprocessing approach is suitable for face authentication: results are comparable with, or even better than those obtained using the state-of-the-art preprocessing algorithm proposed in [9]. Moreover, the LBP representation has several other advantages: there is no need for hyperparameter selection, hence avoiding extensive experiments. Indeed, the normalization results were obtained with different values of λ, depending on the database and on the classifier (a recent study [20] shows that better results are obtained if the normalization parameter is tuned for each probe image) whereas the LBP preprocessing was always the same. Furthermore, the LBP operator is simpler and faster than the illumination normalization process.

5 Conclusion

In this paper, we proposed an original preprocessing technique based on LBP for illumination robust face authentication. The LBP operator provides a texture representation of the face which improves the performances of two different classifiers (holistic and feature-based) as compared to histogram equalization. Moreover, obtained results are comparable when using the state-of-the-art preprocessing algorithm [9] on the same databases [20], and removes the need for parameter selection.

References

[1] T. Ahonen, A. Hadid, and M. Pietikäinen. Face recognition with local binary patterns. In *European Conference on Computer Vision (ECCV)*, 2004.

[2] R. Basri and D.W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(2):218–233, 2003.

[3] Peter N. Belhumeur, Joao Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[4] F. Cardinaux. Local Features and 1D-HMMs for Fast and Robust Face Authentication. IDIAP-RR 17, IDIAP, 2005.

[5] F. Cardinaux, C. Sanderson, and S. Bengio. Face Verification Using Adapted Generative Models. In *IEEE Automatic Face and Gesture Recognition (AFGR)*, 2004.

[6] F. Cardinaux, C. Sanderson, and S. Marcel. Comparison of MLP and GMM classifiers for face verification on XM2VTS. In *Audio- and Video-based Biometric Person Authentication, AVBPA'03*, 2003.

[7] E. Bailly-Baillière et al. The banca database and evaluation protocol. In *Audio- and Video-based Biometric Person Authentication, AVBPA'03*, 2003.

[8] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.

[9] R. Gross and V. Brajovic. An image preprocessing algorithm for illumination invariant face recognition. In *Audio- and Video-Based Biometric Person Authentication, AVBPA'03*, 2003.

[10] M. Heikkilä, M. Pietikäinen, and J. Heikkilä. A texture-based method for detecting moving objects. In *British Machine Vision Conference (BMVC)*, 2004.

[11] G. Heusch, F. Cardinaux, and S. Marcel. Efficient diffusion-based illumination normalization for face verification. IDIAP RR 46, IDIAP, 2005.

[12] H. Jin, Q. Liu, H. Lu, and X. Tong. Face detection using improved LBP under bayesian framework. In *Third International Conference on Image and Graphics (ICIG), Hong Kong, China*, pages 306–309, 2004.

[13] J. Kittler, Y. Li, and J. Matas. On matching scores for lda-based face verification. In *British Machine Vision Conference (BMVC)*, 2000.

[14] S. Marcel. A symmetric transformation for lda-based face verification. In *IEEE Automatic Face and Gesture Recognition (AFGR)*, 2004.

[15] K. Messer, J. Matas, J. Kittler, J. Lüttin, and G. Maitre. XM2VTSDB: The extended M2VTS database. In *Audio- and Video-based Biometric Person Authentication, AVBPA'99*, 1999.

[16] A. Nefian. Embedded bayesian networks for face recognition. In *IEEE International Conference on Multimedia, ICME*, 2002.

[17] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996.

[18] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[19] J. Short, J. Kittler, and K. Messer. A comparison of photometric normalisation algorithms for face verification. In *IEEE Automatic Face and Gesture Recognition (AFGR)*, 2004.

[20] J. Short, J. Kittler, and K. Messer. Photometric normalisation for face verification. In *Audio- and Video-based Biometric Person Authentication, AVBPA'05*, 2005.

[21] V. Takala, T. Ahonen, and M. Pietikäinen. Block-based methods for image retrieval using local binary patterns. In *Proc. 14th Scandinavian Conference on Image Analysis (SCIA), Joensuu, Finland*, pages 882–891, 2005.

[22] M. Turtinen, M. Pietikäinen, and O. Silven. Visual characterization of paper using isomap and local binary patterns. In *Conference on Machine Vision Applications (MVA)*, 2005.

[23] G. Zhang, X. Huang, S.Z. Li, Y. Wang, and X. Wu. Boosting local binary pattern (LBP)-based face recognition. In *Chinese Conference on Biometric Recognition, SINOBIOMETRICS 2004*, 2004.

[24] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003.