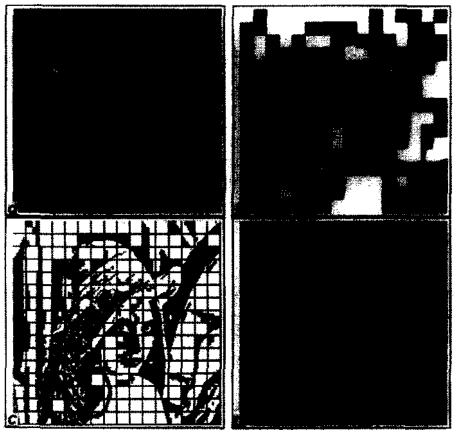

Fig. 1 *Illustration of coding principle*

idea is to expend code-bits at all instances where the marginal gain is greater than μ distortion-units/bit. For example, assume some part of the image can be encoded in two different ways. Either it costs R bits and the distortion is D, or it costs $R + \Delta R$ bits and the distortion is $D - \Delta D$. Then, we will choose the more expensive alternative if and only if $\Delta D/\Delta R \geq \mu$ distortion-units/bit.

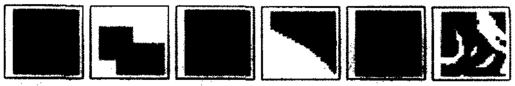

Fig. 2 *Example of three enlarged source-blocks and their optimised fragmentation*

The fragmentation is described in high resolution only if the two fragments have very different statistics. Sites in a fragment do not need to be connected.

As was hinted at above, the fragmentation is encoded by a bintree. By this method, the fragmentation can be described in high resolution near sharp edges, and in low resolution at places where the exact shape of a fragment has a minor impact on the distortion (see Fig. 2). To find the optimal bintree, we start with a maximum tree, and prune this tree by cutting off all branches where the marginal gain is less than μ distortion-units/bit.

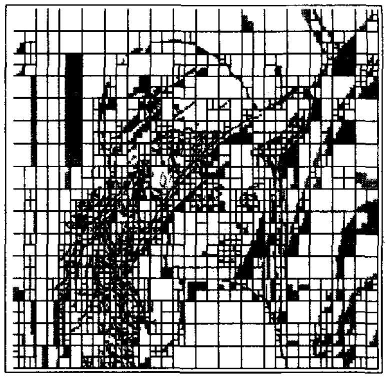

Fig. 3 *Fragmentation of Lena.512 image*

In each fragment, the six transform components are uniformly quantised and Huffman encoded. It can be shown [1] that the optimised quantisation step is given by

$$\Delta = \sqrt{\frac{12\mu}{2\ln 2}}$$

Results: Fig. 3 illustrates a coder which is using both fragmented blocks and variable block size. The fragmentation, the variable block size, and the quantisation of the transform components are all controlled by the optimality principle. The test image is Lena.512 (Fig. 4). The rate is 0.23bit/pixel, and the peak signal to noise ratio (PSNR) = 30.9dB. As a comparison, JPEG attains *PSNR* = 29.8dB at the same bit rate.

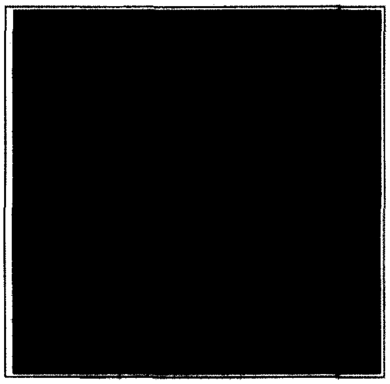

Fig. 4 *Lena.512 image*

Rate = 0.23bit/pixel (i.e. compression ratio = 35); *PSNR* = 30.9dB

© IEE 1995 *24 March 1995*
Electronics Letters Online No: 19950583

R. Nohre (*Division of Information Theory, Department of Electrical Engineering, Linköping University, 58183 Linköping, Sweden*)

References

1 NOHRE, R.: 'Image coding by fragmented blocks' Technical Report, 1995, Linköping University, Sweden (LiTH-ISY-R-1734)

2 KUNT M., BENARD M. and LEONARDI R.: 'Recent results in high-compression image coding', *IEEE Trans.*, **CAS-34**, (11), pp. 1306–1336.

3 FORCHHEIMER R. and KRONANDER T.: 'Image coding – from wave forms to animation', *IEEE Trans. Acoust. Speech Signal Process.*, 1989, pp. 2008–2023.

4 BIGGAR M.J., MORRIS O.J. and CONSTANTINIDES A.G.: 'Segmented-image coding: performance comparison with the discrete cosine transform', *IEE Proc. F*, **135**, (2), pp. 121–132.

5 NAISOPOULOS P., WARD R.K. and MORSE D.J.: 'Adaptive compression coding', *IEEE Trans.*, **COM-39**, (8), pp. 1245–1254.

6 DELP E.J. and METICHELL O.R.: 'Image compression using block truncation coding', *IEEE Trans.*, **COM-37**, (9), pp. 1335–1342.

7 WU H.-S., KING R.A. and KITNEY R.I.: 'Improving the performance of the quadtree based image approximation via generalised DCT', *Electron. Lett.*, **29**, (10), pp. 887–888.

System approach to disparity estimation

S. Panis, M. Ziegler and J.P. Cosmas

Indexing terms: Stereo image processing, Dynamic programming

A system approach to disparity estimation using dynamic programming is presented. The four step system can calculate a dense correspondence map between a stereo pair with parallel or nonparallel camera geometry. Results are presented with CCIR 601 format stereo images.

Introduction: Binocular stereo is the process of obtaining depth information from a pair of left and right camera images [1]. Surveys on stereo matching techniques can be found in [2, 3]. Many approaches are system oriented like the approach described here.

The system consists of four parts (Fig. 1):

(i) global modelling of the luminance difference between the stereo images

(ii) determination of useful disparity search range along each scan line

(iii) disparity estimation based on 3-D dynamic programming

(iv) disparity map postprocessing

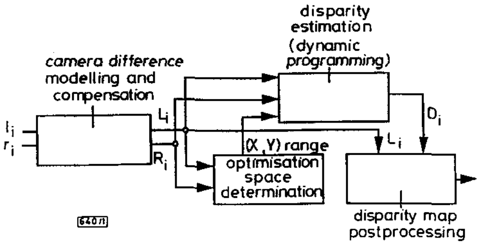

Fig. 1 *Disparity estimation system*

(i) Global modelling of luminance difference between left and right images: In a stereo pair of images, there may be an imbalance in the luminance levels between the left and the right images, caused by different camera configurations and characteristics. Therefore, statistical modelling and compensation of the luminance difference take place in this system as a preprocessing step, based on a method described in [4].

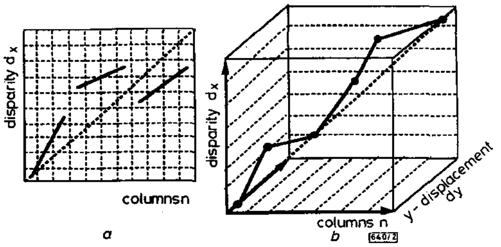

Fig. 2 *Optimum path finding*

a 2-D optimisation space
b 3-D optimisation space

(ii) Determination of useful disparity search range along scan lines: The number of paths that have to be searched by dynamic programming is $(d_x{\cdot}d_y)^n$ where n is the column axis, d_x is the disparity axis and d_y is the y-displacement axis of the optimisation space (Fig. 2*b*). As the range of disparity and y-displacement increase, the execution time grows and hence the danger of considering paths that may lead to wrong results increases. Therefore, a preprocessing stage calculates the maximum and minimum values of disparity and y-displacement, which are subsequently used to restrict the size of the dynamic programming optimisation space. The calculation is performaed in three steps. First, a large-block-matching routine gives an estimate of the range of displacement vectors. Secondly, a block-based median filtering smoothens the disparity map and finally, scanning each scan line, the minimum and maximum vectors are calculated.

(iii) Disparity estimation based on 3-D dynamic programming: In [1, 5], two dynamic programming based algorithms for disparity estimation are described. The algorithms work with ideal stereo images. Therefore, the dynamic programming optimisation space is two-dimensional (Fig. 2*a*) because optimisation is performed only on disparity. In the case of nonparallel geometry or image misalignment, the algorithms fail because the y-displacement is nonzero. In this work, a 3-D optimisation space is used because y-displacement is also considered (Fig. 2*b*). This way, calculation of epipolar geometry is not required and a parameter free system is developed that is able to function with imperfect stereo images. The total cost function for dynamic programming is a combination of a 'normalised matching cost' and a 'disparity jump cost'. The normalised matching cost (NMC) is determined by the quality of matching whereas the disparity jump cost is a regularisation factor which enforces disparity to be smooth along a uniform surface (monotonicity constraint [1]). Where disparity fails to be smooth, the area is classified as occluded. The disparity jump cost is defined, as in [1], by

$$f(x) = \mu\sqrt{x} + \varepsilon|x| \qquad x \geq 0 \tag{1}$$

where μ = 0.3 and ε = 0.15; empirically determined in [1]. The total cost $G_n(j, i)$ is calculated with the recursive function:

$$\begin{aligned} G_n(j,i) &= NMC_n(j,i) + f(x) + G_n(j_{best}, i_{best}) && n > 1 \\ G_n(j,i) &= NMC_n(j,i) && n = 1 \end{aligned} \tag{2}$$

where x is the disparity jump between the current candidate disparity and the disparity of the previous column, n is the current column number, j is the candidate disparity, i is the candidate y-displacement and (j_{best}, i_{best}) is the best predecessor vector.

(iv) Disparity map postprocessing: Postprocessing is designed to smoothen the disparity map by correlating each disparity value to its neighbouring values and filling up gaps where disparity failed to be calculated. Postprocessing of disparity is divided into three parts:

(*a*) median filtering with 3 × 3 window (neighbourhood correlation)

(*b*) median filtering with 1 × 5 window (vertical dispariry correlation)

(*c*) interpolation (gap filling)

Fig. 3 *Original left even field of first frame of 'Aqua' image sequence*

Fig. 4 *Disparity map for stereo pair 'Aqua'*

Fig. 5 *Disparity map after postprocessing for stereo pair 'Aqua'*

Results: The stereo image pair 'Aqua' [6] (Fig. 3), has disparity range in the interval +1 to +34 pixels, calculated from the left to the right images. The image size is 720 × 288 and the calculation time was 8 min on a Sparc 10 workstation. From the equalised disparity map shown in Fig. 4, the depth of each object compared to the other objects can be seen. The objects with a lighter grey value (larger disparity) are closer to the camera whereas those with a darker grey value (smaller disparity) are further away from the camera. Black areas have no disparity value because they were classified as occlusion regions. By visual comparison between the original picture and its disparity map, we can recognise the shapes of the objects. Good examples are the fish, plants and the rock. Fig. 5 shows the postprocessed disparity map. It is smoother and the object definition is clearer. The disparity map of Fig. 5 has been used to reconstruct the left image from the right image. The reconstructed image is visually good and has a *PSNR* = 29dB, which shows that the disparity fidelity is quite high. Block-based disparity gave a *PSNR* = 23dB.

Acknowledgment: This work was carried out within the framework of RACE (Research on Advanced Communications in Europe) under contract R2045 'DISTIMA' (DIgital STereoscopic IMaging and Applications).

© IEE 1995 *20 March 1995*
Electronics Letters Online No: 19950617

S. Panis and J.P. Cosmas (*Electronic Engineering Dept., Queen Mary and Westfield College, Mile End Road, London E1 4NS, United Kingdom*)

M. Ziegler (*Siemens AG, ZFE T SN 22, 81730 Munich, Germany*)

S. Panis: also with Siemens AG, Munich, Germany

References

1 GEIGER, D., LANDENDORF, B., and YUILLE, A.: 'Occlusions and binocular stereo', *Int. J. Comput. Vis.*, 1993,

2 BARNARD, S.T., and FISCHLER, M.A.: 'Computational stereo', *Computing Surveys*, 1982, **14**, (4), pp. 553–572

3 DHOND, U.R., and AGGARWAL, J.K.: 'Structure from stereo-A review', *IEEE Trans. Syst. Man Cybern.*, 1989, **19**, (6), pp. 1489–1510

4 'Balance compensation for stereoscopic sequences' RACE DISTIMA Doc. 45/TUD/WP3.2/DN/C/27.792/1

5 COX, I.J., HINGORANI, S., MAGGS, B.M., RAO, S.B.: 'Stereo without regularization'. NEC Research Institute Report, 21 October 1992

6 Stereoscopic test sequences (CCIR 601 format) shot within RACE DISTIMA

Searching for the best linear approximation of DES-like cryptosystems

L. Buttyan and I. Vajda

Indexing terms: Cryptography, Graph theory, Viterbi decoding

The authors show that the problem of searching for the best characteristic in linear cryptanalysis is equivalent to searching for the maximal weight path in a directed graph. Under certain restrictions, the best characteristic can be easily obtained even on a personal computer.

Introduction: The design of information systems should take into account potential security problems. Cryptography provides a strong information processing technique for algorithmic protection of information. Differential cryptanalysis [1] and linear cryptanalysis [2] are the most effective methods to date for attacking DES-like cryptosystems. Differential and linear cryptanalyses are very similar on a structural level as shown in [3, 4]. Both methods rely on the existence of an *r* round characteristic having high probability, where *r* is the number of rounds of the cryptosystem. Therefore, searching for the best characteristic is a very important part of these attacks. However no practical algorithm has yet been published which finds the best characteristic for the full number of rounds without any restrictions. We introduce an abstract model for the searching problem based on searching for the maximal weight path in a directed graph. In this Letter, we focus on linear cryptanalysis, but our results can be adopted for differential cryptanalysis, owing to their duality [4].

The aim of linear cryptanalysis is to find out the key (or a part of the key) of the cryptosystem with the help of an effective linear approximation

$$\Omega \cdot P + \Theta \cdot C + \sum_{i=1}^{r} \Phi_i \cdot K_i = 0 \qquad (1)$$

of the cipher and a large number of plaintext-ciphertext pairs. Here, *P* and *C* denote the plaintext and the ciphertext blocks, respectively, K_i stands for the subkey of the *i*th round, Ω, Θ and Φ_i are the mask vectors. Furthermore, '·' and '+' stand for the binary scalar product and XOR operation, respectively. The effectiveness of eqn. 1 depends on how well it agrees with the probability Q'. A larger value of $|Q' - 1/2|$ provides a more effective linear approximation of the cipher.

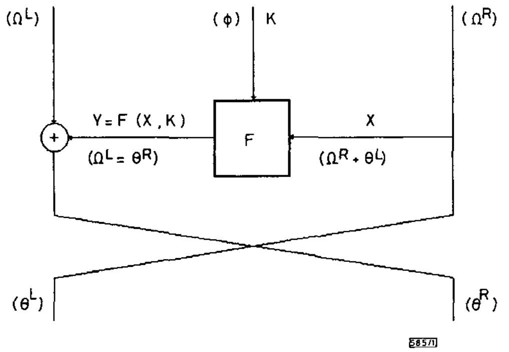

Fig. 1 *Structure of 1 round and the masking vectors*

Characteristics: We define an *r* round characteristic to describe a linear approximation of an *r* round cipher. Our definition is not exactly the same as that of Biham [3], but it is better suited to our purposes.

Definition 1: A 1 round characteristic is a tuple $\chi = (\Omega, \Theta, \Phi, q')$, where $\Omega = (\Omega^L, \Omega^R)$, $\Theta = (\Theta^L, \Theta^R)$ and $\Omega^L = \Theta^R$. q' is the probability that the linear approximation

$$(\Omega^R + \Theta^L) \cdot X + \Omega^L \cdot Y + \Phi \cdot K = 0 \qquad (2)$$

of round function *F* holds, where $Y = F(X, K)$ (see Fig. 1).
Definition 2: An *r* round characteristic ($r > 1$) is a series of *r* 1 round characteristics $\chi_1, \chi_2, \ldots, \chi_r$, where $\chi_i = (\Omega_i, \Theta_i, \Phi_i, q'_i = 1/2 + q_i)$ such that $\Omega_i = \Theta_{i-1}$ for $i = 2, 3, \ldots, r$. Thus an *r* round characteristic determines *r* linear approximations

$$(\Omega_i^R + \Theta_i^L) \cdot X_i + \Omega_i^L \cdot Y_i + \Phi_i \cdot K_i = 0 \qquad (3)$$

with probability $q'_i = 1/2 + q_i$, $0 < i < r + 1$.

In DES-like cryptosystems,

$$Y_i = X_{i-1} + X_{i+1} \qquad 0 < i < r + 1 \qquad (4)$$

where $Y_i = F(X_i, K_i)$, and $(X_0, X_1) = (P^L, P^R) = P$ and $(X_r, X_{r+1}) = (C^L, C^R) = C$ are the plaintext and ciphertext blocks, respectively. We omit the swapping of the two halves of *C*, because it does not have any cryptographic significance.

Since $\Omega_i^L = \Theta_i^R$, $i = 1, 2, \ldots, r$ by definition 1 and $\Omega_i = \Theta_{i-1}$, $i = 2, 3, \ldots, r$ by definition 2, $\Omega_{i-1}{}^L + \Omega_{i+1}{}^L + (\Omega_i^R + \Theta_i^L) = 0$ for $i = 2, 3, \ldots, r-1$. Summing up the *r* expressions of eqn. 3 we get

$$\Omega_i \cdot P + \Theta_r \cdot C + \sum_{i=1}^{r} \Phi_i \cdot K_i = 0 \qquad (5)$$

The probability of eqn. 5 is $Q' = 1/2 + 2^{r-1} q_1 q_2 \ldots q_r$ according to the piling-up lemma in [2]. So the problem is to find an *r* round characteristic for which $\Pi |q_i|$ is maximal.

Graph representation and searching algorithm: We can represent all the 1 round characteristics with a two-stage directed graph. The nodes of the first stage and the second stage of the graph represent the possible values of Ω and Θ, respectively. An edge leads from a node of the first stage, say from ω, to a node of the second stage, say to θ, if there exists a 1 round characteristic $\chi = (\omega, \theta, 1/2+q)$, for which $q \neq 0$. The weight of this edge is $\log|q|$.