Document No.: 88
March 1986

CCITT STUDY GROUP XV
Working Party on Visual Telephony
Specialists Group on Coding for Visual Telephony

SOURCE: Federal Republic of Germany

TITLE: STATISTICAL PROPERTIES AND ENCODING OF DISPLACEMENT VECTORS OBTAINED BY A BLOCK-BASED METHOD

1. Scope of this contribution

If in a motion compensating coder for video sequences displacement vectors are estimated from data, that are not available at the receiver, these vectors have to be transmitted. This contibution addresses the question, how displacement vectors can be encoded efficiently on a block-by-block basis, i.e. one vector per NxN-block is transmitted. For the videoconference format of 352x288 pels per field, for each field up to 1584 vectors have to be transmitted for 8x8-blocks, or up to 396 vectors for 16x16-blocks. On the other hand, the bit budget for transmission rates of 384 kbit/s and below allows 25000 bits or less for the transmission of a single field, including displacement vectors.

In this contribution, we consider three approaches to code the displacement vector:

A) Fixed wordlength coding of the displacement vector

A fixed number of bits is used for each displacement vector. If there are M different possible displacement values, at least $\log_2 M$ bits are neccessary.

B) Variable wordlength coding of the displacement vector

A Huffman code is used for the displacement vector.

C) Variable wordlength coding of the differences of horizontally adjacent displacement vectors

A Huffman code is used for the vector-valued difference between horizontally adjacent displacement vectors: the displacement value of the previous block is used as prediction of the current displacement vector. Koga et. al. [1] report a reduction of the data rate for the displacement vector by 40% with this approach in conjunction with a specific displacement estimation method.

2. Statistical properties of the displacement vector

The three approaches A), B), and C) have been compared statistically for different sequences and several block-based displacement estimation methods. Generally it can be observed that for displacement vectors other than the static-background-vector (0,0) the x- and y-displacement are statistically independent.

As an example of their statistical behaviour Fig. 1 shows the entropies of the x- and y-displacement measured for the sequence "Split screen/Trevor" at a field frequency of 12.5 Hz with Jain and Jain's 2-D-logarithmic search method [2] for blocks of 16x16 and a maximum displacement of 7. The entropy measurement considers all blocks of each field including static background and corresponds to the approaches A), B), and C), as described above. Integer-pel-accuracy of the displacement vector is used, which is sufficient for most scenes, not because of insufficient displacement estimation methods, but due to the inaccuracy of the image model "translatory displacement" [3]. With approach A) 2x4 bits (lower bound: $2 \cdot \log(15)$) are required for each displacement vector.

Approach B), applied to the components of the displacement vector independently, reduces the required amount of information to approximately 50% compared to fixed length encoding. The entropy for the overall picture, $H_{overall}$ is determined by the entropy within moving parts, H_{moving} and the percentage of moving blocks, P_{moving}, according to

$$H_{overall} = P_{moving} \cdot H_{moving} + \mathcal{H}(P_{moving})$$

where

$$\mathcal{H}(X) = - X \cdot \log_2(X) - (1-X) \cdot \log_2(1-X)$$

Most of the gain of variable length coding of the displacement vector (approach B) over fixed length coding (approach A), as suggested by Fig. 1, is due to the high percentage of displacement vector (0,0) in the static background (Fig. 2). If the entropy of the displacement is considered in moving areas only, this gain is typically smaller (Fig. 3). Thus, a reasonable, robust, although suboptimum code for the case considered could consist of one bit indicating whether the vector is (0,0) or not, and, in the case of non-zero vector, two 4-bit-words for the vector components, resulting in a typical rate of 4 bit/block for the displacement vectors of Split screen/Trevor in the average.

Approach C) does not result in further data reduction, but usually requires slightly more bits for the displacement vector than B) (Figs. 1 and 2). This is due to the fact, that Jain and Jain's block matching method yields a good vector for prediction, which, however, does not neccessarily correspond to the true motion in the scene.

With displacement estimation methods, that take into account the scene contents of more than a single block, such as [1], or such as object matching approaches [4], it can be observed, that the difference of adjacent displacement vectors has a considerably smaller entropy than the vectors themselves. The impact of the displacement estimation method on the behaviour of the vectors can be as extreme as illustrated in Fig. 4 for a sequence that contains high motion. Consequently, the coding standard for the transmission of displacement vectors on a block-by-block-basis should allow a further reduction of the displacement vector transmission rate by advanced methods. This can be achieved by one bit per transmitted field, that indicates whether displacement vector amplitude values or difference values are transmitted. A variable length code for the displacement differences has to be designed based on model assumptions, as the ultimate displacement estimation method is not yet known.

Observations very similar to the ones described above have been made for other sequences and for other blocksizes.

3. Conclusions

A coding standard for the transmission of displacement vectors on a block-by-block basis should allow both a transmission of vector amplitude values and vector difference values with variable length codes. Except for the case of the static background vector (0,0), an independent encoding of the vector components is adequate. For vector amplitude encoding a large fraction of the gain by variable length encoding is due to the high probability of the vector (0,0) in the static background. A variable length code for the displacement differences has yet to be designed.

4. References

[1] T. Koga, A. Hirano, K. Iinuma, Y. Iijima, T. Ishiguro, "A 1.5 Mb/s Interframe Codec with Motion-Compensation", IEEE International Conference on Communications 1983, Boston, pp. D8.7.1 - D8.7.5.

[2] J. R. Jain, A. K. Jain, "Displacement Measurement and Its Application in Interframe Image Coding", IEEE Trans. on Communications, Vol. COM-29, No. 12, December 1981, pp. 1799 - 1808.

[3] B. Girod, T. Micke, "Efficiency of motion-compensating prediction in a generalized hybrid coding scheme", Picture Coding Symposium 1986, Tokyo.

[4] G. Kummerfeldt, F. May, W. Wolf, "Coding Television Signals at 320 and 64 kbit/s", 2nd International Technical Symposium on Optical and Electro-Optical Science and Engineering, SPIE Conf. B594 'Image Coding', Cannes, France, Dec. 1985.

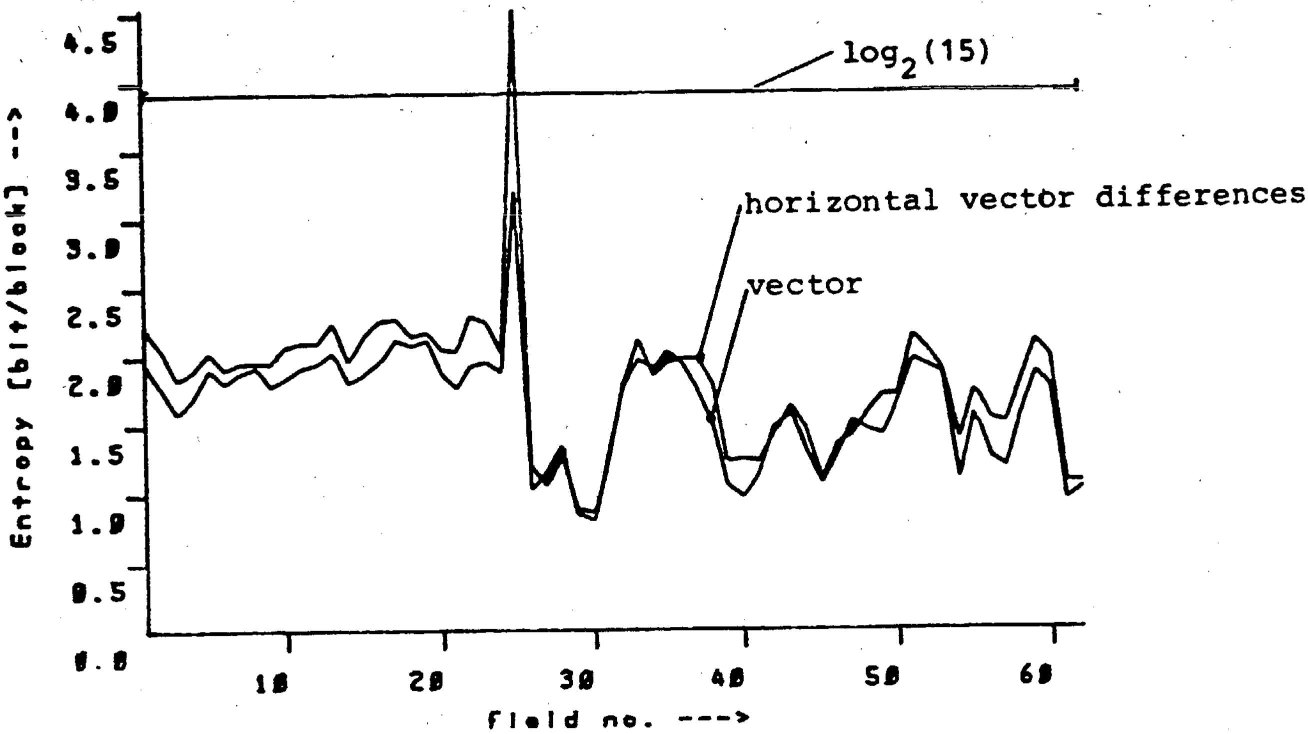


Fig. 1a - Entropy of x-displacement measured with Jain and Jain's method for sequence Split screen/Trevor

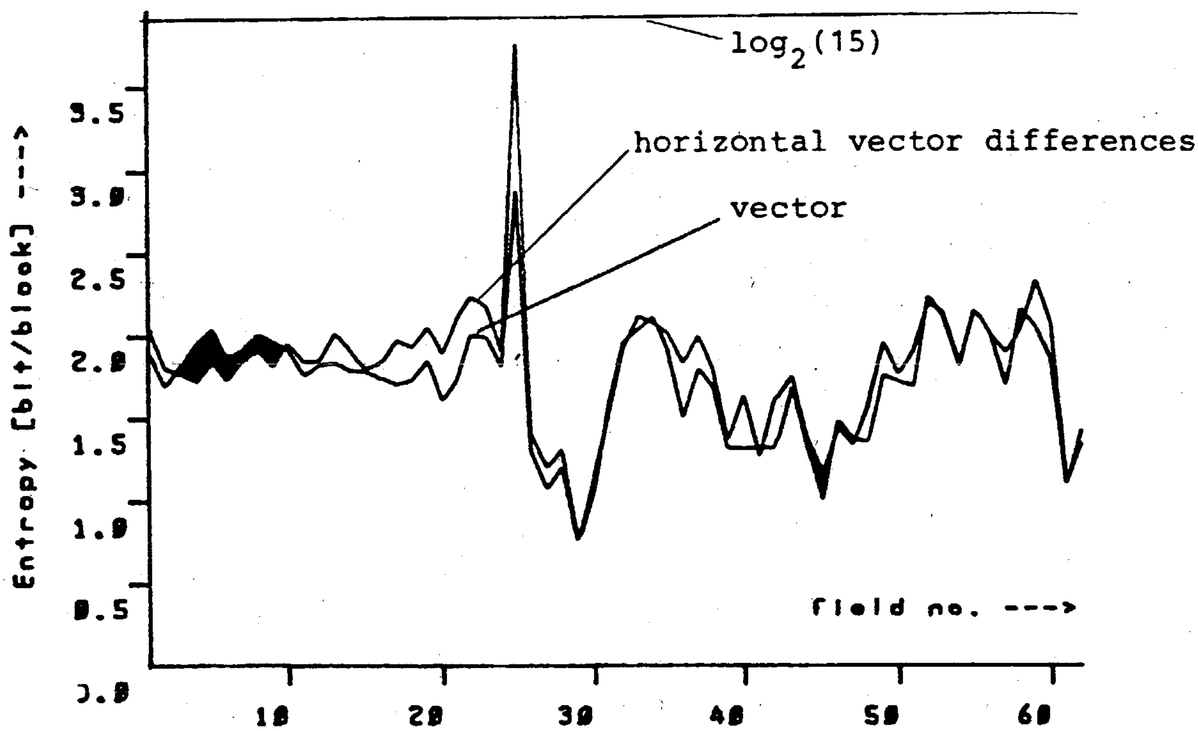


Fig. 1b - Entropy of y-displacement measured with Jain and Jain's method for sequence Split screen/Trevor

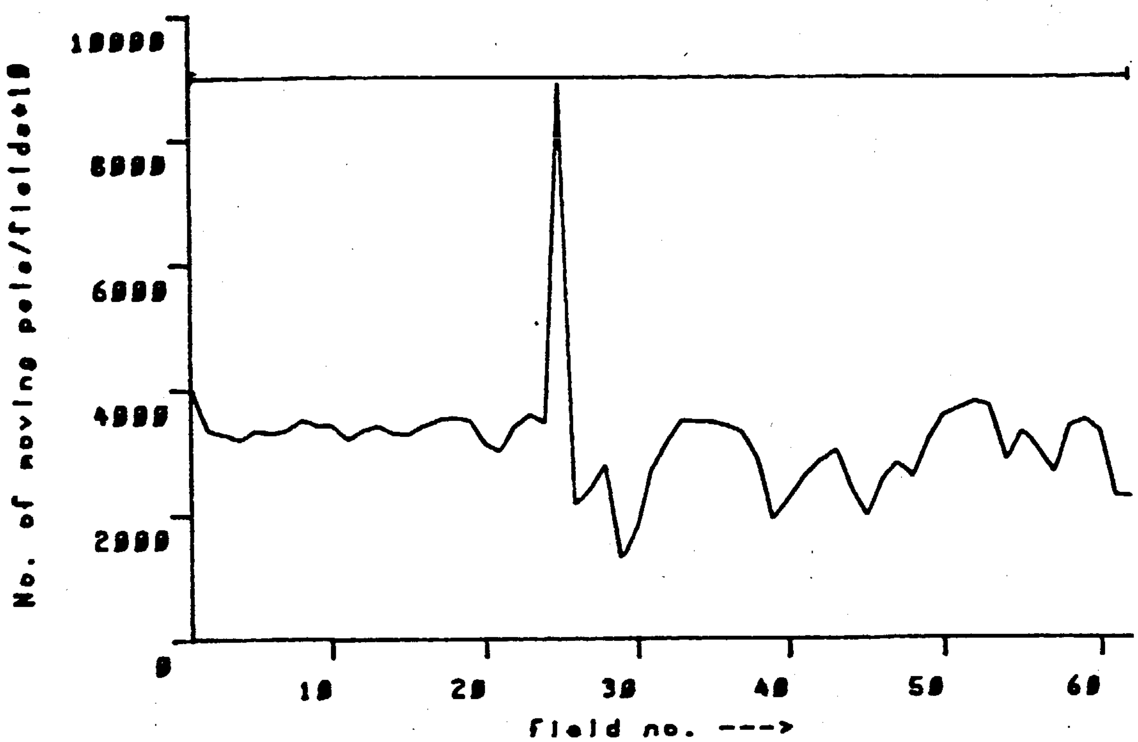


Fig. 2 - Number of moving pels per field of sequence Split screen/Trevor

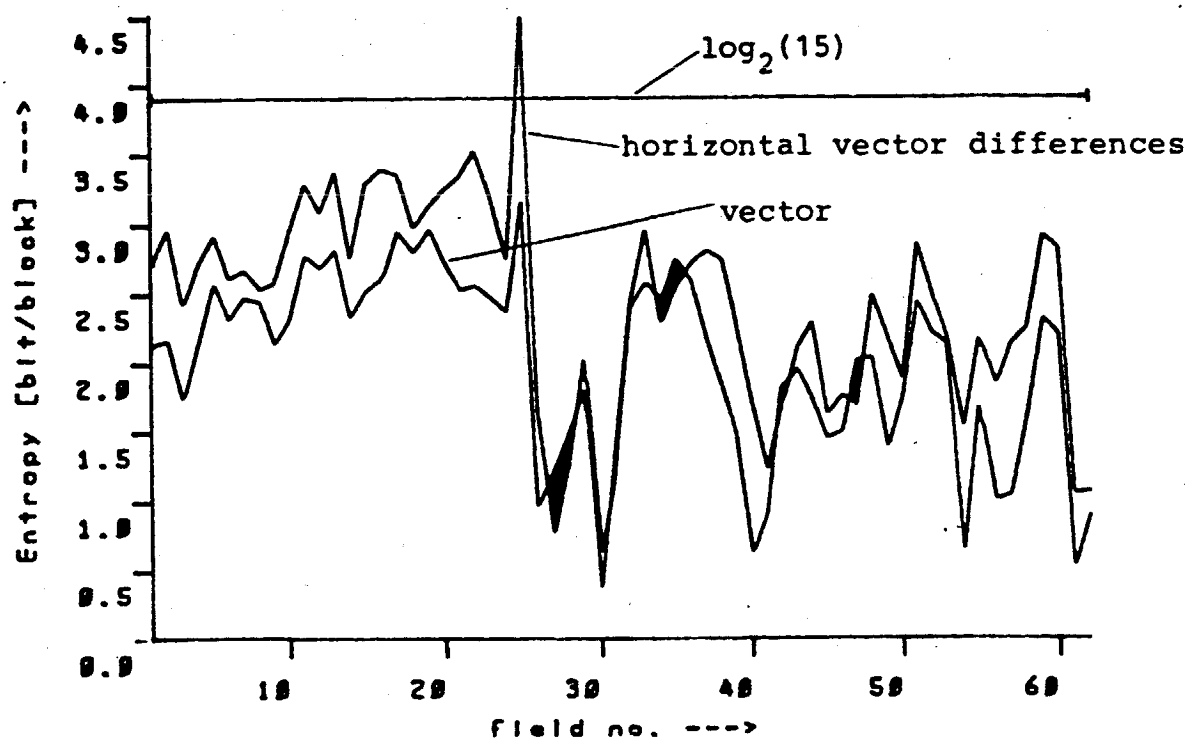


Fig. 3a - Entropy of x-displacement in moving parts measured with Jain and Jain's method for sequence Split screen/Trevor

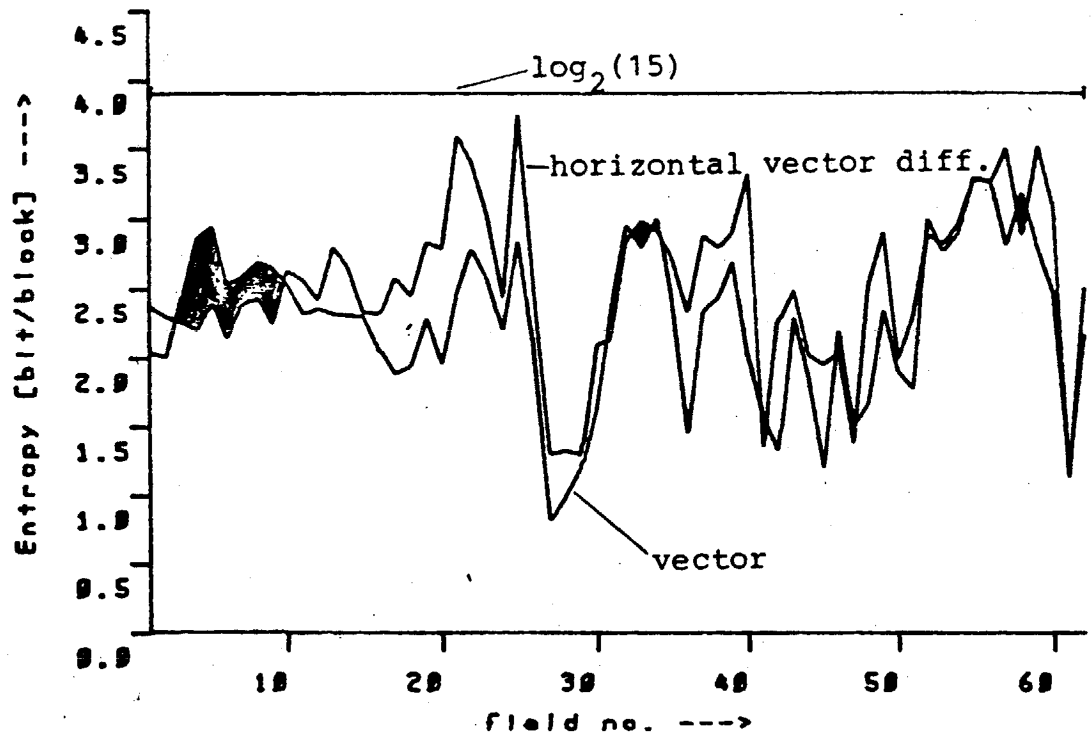


Fig. 3b - Entropy of y-displacement in moving parts measured with Jain and Jain's method for sequence Split screen/Trevor

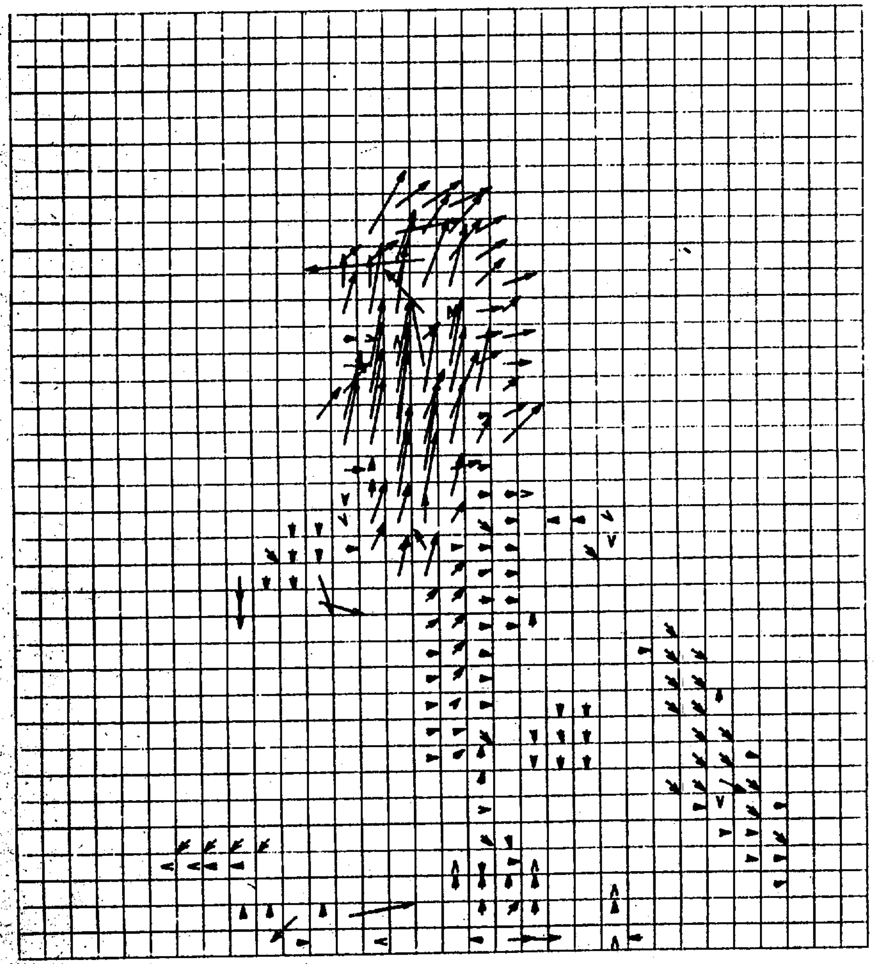
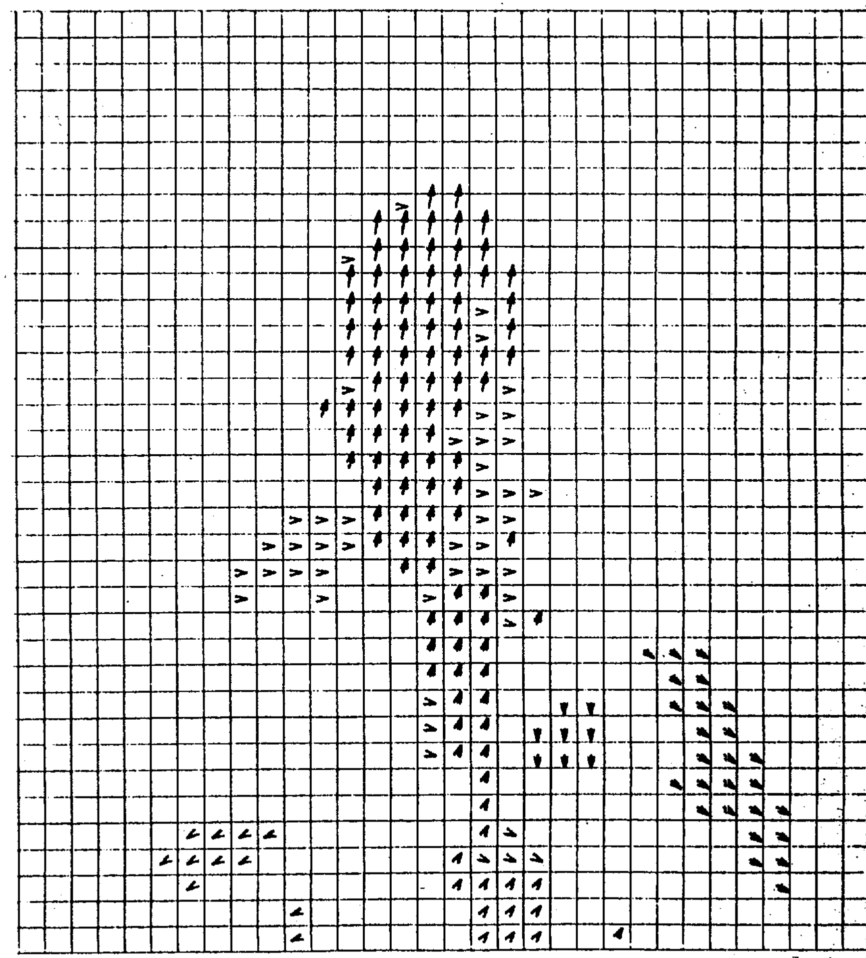


Fig. 4 Block matching (8*8) displacement vector diagram for 'Trevor'



Object matching (8*8) displacement vector diagram for 'Trevor'