MVA '92 IAPR Workshop on Machine Vision Applications Dec. 7-9,1992, Tokyo

Interpretation of Optical Flow through Neural Network Learning

Minami MIYAUCHI, Masatoshi SEKI,[†] Akira WATANABE and Arata MIYAUCHI

School of Management and Informatics,
SANNO College
1573 Kami-kasuya, Isehara,
Kanagawa 259-11 Japan
miyauchi@miyau.mi.sanno.ac.jp

Department of Electronics and Communication
Musashi Institute of Technology
1-28-1 Tamazutsumi, Setagaya-ku,
Tokyo 158 Japan
miyauchi@elcom.musashi-tech.ac.jp

Abstract In computer vision, the interpretation of optical flow (motion vector field calculated from images) and estimation of motion are important tasks. This study proposes a motion interpretation network which enables optical flow (OF) interpretation and describes motions on a plane through the use of a neural network with complex back propagation learning. Furthermore, an OF normalization network for optical flow normalization is proposed for the interpretation of diverse flow patterns, such as real image optical flow. Using test patterns and real image optical flow, the generalization capacity of proposed network is investigated. And the ability is confirmed experimentally.

1 Introduction

In computer vision, the interpretation of optical flow [1] (motion vector field calculated from images) and estimation of motion are important tasks [2]. This study proposes a motion interpretation network which enables optical flow (OF) interpretation and describes motions on a plane through the use of a neural network with complex back propagation learning. Furthermore, an OF normalization network for optical flow normalization is proposed for the interpretation of diverse flow patterns, such as real image optical flow.

Methods for estimating motion from optical flow include a method that obtains the optimum solution by using several flow vectors to solve equations [3], [4], [5]. However, this method is time consuming and prone to noise, and solutions are for actual images cannot easily be obtained.

Neural networks are frequently utilized in pattern translation and are far less affected by noise [6], [7]. The calculation time required after learning is short, and the network are suitable for interpretation of motion. In addition, the networks proposed in this paper utilize complex BP and thus can naturally accommodate optical flow, a two-dimensional vector, as a complex number.

2 Complex Back Propagation Learning

Complex back propagation learning has been developed by Nitta and Furuya [8], who expanded the weight of connection and the threshold of each unit in conventional neural networks to complex numbers. It is shown as an effective method for graphic conversion.

3 Motion Interpretation Network

Figure 1 shows the architecture of the motion interpretation network proposed in this study. In the flow vectors to be fed to the neural network, it is assumed that motion develops all over the frame, centering around the center of the frame.

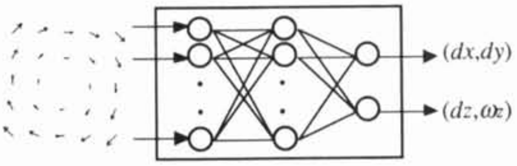


Figure 1 Motion interpretation network

- **Input layer**

  Units corresponding to each vector of optical flow are arranged two dimensionally. Inputs to each unit are of complex numbers corresponding to motion vectors of the inputs. Units of 25 ($= 5 \times 5$) was used in the computer simulation.

- **Output layer**

  Two complex output units corresponding to the displacement components parallel to the frame (dx, dy), expansion and contraction component dz and rotational component parallel to the frame ω_z (dz, ω_z) are available.

- **Hidden layer**

  There is only one hidden layer, and 16 units are used.

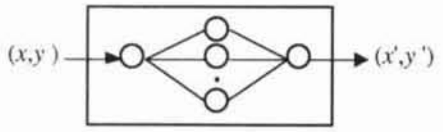
4 Motion Interpretation Networks with Normalization Capacity

Motion interpretation networks interpret the optical flow in an entire frame. However, in reality, optical flow frequently cannot be obtained for the entire frame. Hence, OF normalization networks utilize the graphic conversion networks proposed by Nitta and Furuya to normalize sparce optical flow, partially defined optical flow and optical flow of arbitrary size and shape to $n \times n$.

In OF normalization networks, a complex function that translates a point on a two-dimensional plane to another point can be estimated by supplying a point before translation and a point after translation as learning data.

By having the network learn the starting point of each vector of optical flow as a point before conversion and the terminal point as a point after conversion, it is possible to have the neural network estimate a function f, to describe the optical flow. Here the network is expected to output a value of the complex function f, at all locations on the frame. By providing points aligned in an $n \times n$ lattice format to this network as starting points, and obtaining the final points from each starting point, a normalized $n \times n$ optical flow can be created. Figure 2 shows the architecture of OF normalization network.

[†]Now with Toshiba medical engineering Co.



Function f

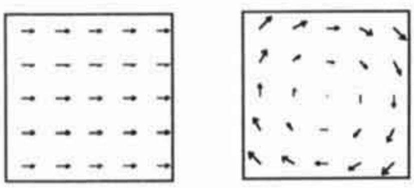
Figure 2 OF normalization network

5 Experimental

A test pattern consisting of 25 ($= 5 \times 5$) vectors was provided to a motion interpretation network which had learned basic motions, so that the generalization capacity of the motion interpretation network could be studied. The teacher pattern and test pattern provided motion components, and these were arranged so that motion developed all over the frame, centering around the center of the frame.

The number of types of tercher patterns used was 25, and these patterns contained a single motion component. Then 47 types of test patterns were provided.

Table 1 shows the motion vectors of teacher pattern supplied and Figure 3 shows an example of teacher pattern. A motion vector of test pattern and motion interpretation network output corresponding to the pattern are given in Tables 2 a) and 2 b), respectively. Patterns 1 to 24 are patterns for investigating generalization capacity with respect to unknown speeds; patterns 25 to 35 are patterns for investigating the generalization capacity for a combination of multiple numbers.



a) $dx = 1$ b) $\omega_z = 1$

Figure 3 Exanples of teacher patterns

The results indicate that although ununiformity was present, the mean square error was 0.0016 on the average, with a maximum of about 0.055. Motion interpretation networks that have learned basic mation are thought to have the capacity to generalize unknown patterns.

Next, examples of experinental results on normalization capacity are shown. In the experiments, patterns in which optical flow was not obtained for the entire frame (test patterns 36–47) were created, then normalized by the OF normalization network and interpreted by the motion interpretation network. Figure 5 shows an input optical flow, the same optical flow after normalization, and the results of interpretation for each of them. In Figure 5, an input optical flow of test pattern is shown above, its normalized pattern is shown below, resuls of interpretation and provided motion vector are indicated as $(dx,\ dy)(dz,\ \omega_z)$ at the bottom.

Figure 6 shows an input optical flow with aditive noise($\sigma^2 = 0.00067, 0.0053$ and burst noise), the same optical flow after normalization, and the results of interpretation for each of them. This result indicates the interpretation network and the normalization network works well even under the additive noise.

Figure 7(a) shows the first frame of real image sequence(128 pixels $\times$ 128 pixels) used in this experiment and the object moves about 2.5 pixels to the left. Figure 7(b) shows a part of optical flow (10 pixels $\times$ 10 pixels) caluculated from the real image sequence. Figure 8 shows the pattern after normarization. The result of this experiment shows the effectiveness of our networks.

Table 1 Motion vector of teacher patterns

	dx	dy	dz	ωz
1	1	0	0	0
2	0	1	0	0
3	-1	0	0	0
4	0	-1	0	0
5	1	1	0	0
6	-1	1	0	0
7	1	-1	0	0
8	-1	-1	0	0
9	0	0	1	0
10	0	0	-1	0
11	0	0	0	1
12	0	0	0	-1
13	0.5	0	0	0
14	0	0.5	0	0
15	-0.5	0	0	0
16	0	-0.5	0	0
17	0.5	0.5	0	0
18	-0.5	0.5	0	0
19	0.5	-0.5	0	0
20	-0.5	-0.5	0	0
21	0	0	0.5	0
22	0	0	-0.5	0
23	0	0	0	0.5
24	0	0	0	-0.5
25	0	0	0	0

6 Conclusions

Through the use of the method proposed in this paper, it was possible to interpret the optical flow and obtain motion parameters on a plane from optical flow. Equipping the network with normalization capacity further enabled us to obtain motion components from various optical flows. In future studies, we shall attempt to develop the present method to enable extraction of 3D motion.

References

[1] B.K.P. Horn and B.G. Schunck: "Determining Optical Flow", Artificial Intelligence, 17, pp.185–203 (1981)

[2] J.K. Aggarwal and N. Nandhakumar: "On the computation of motion from sequences of images — A review", Proceedings of IEEE, 76, 8, pp.917–935 (1988)

[3] R.Y. Tsai and T.S. Huang: "Uniqueness and Estimation of Theree-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-6, No.1, pp.13–27 (1984)

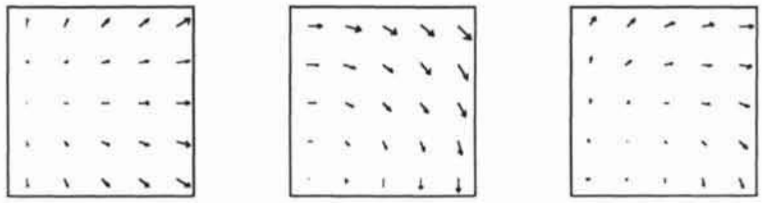
[4] G. Adiv: "Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-7, No.4, pp.384–401 (1985)

[5] M. Subbarao: "Interpretation of Image Flow: A Spatio–Temporal Approach", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-11, No.3, pp.266–278 (1989)

[6] J. Hutchinson, C. Koch, J. Luo and C. Mead: "Computing Motion Using Analog and Binary Resistive Networks", IEEE Computer, 21, 3, pp.52–63 (1988)

[7] K. Imai, K. Gouhara and Y. Hchikawa: "Detection of Position and Size Using Newral Network Learning", Trans. Inst. of Elec., Inform., and Comm. Eng., Vol.J74–D–2, No.6, pp.748–756 (1991) (in japanese)

[8] T. Nitta and T. Furuya: "A Complex Back Propagation Learning", Trans. Information Proccesing Society of Japan, Vol.32, No.10, pp.1319–1329 (1991) (in japanese)



	test pattern 25	test pattern 30	test pattern 31
motion vector of test pattern	$dx = 0.5$ $dy = 0.5$	$dx = 0.5$ $dy = 0.5$ $\omega z = 0.5$	$dx = 0.33$ $dz = 0.33$ $\omega z = 0.33$
output of network	$dx = 0.526$ $dz = 0.552$	$dx = 0.525$ $dy = 0.520$ $\omega z = 0.520$	$dx = 0.367$ $dz = 0.388$ $\omega z = 0.394$

Figure 4 Examples of test patterns

Table 2 Motion vector of test pattern and output of network

	a) motion vector of test pattern				b) output of network				c) MSE
	dx	dy	dz	ωz	dx	dy	dz	ωz	
1	0.75	0	0	0	0.7568	0.0046	0.0096	-0.0031	0.00008459
2	0	0.75	0	0	0.0065	0.7546	0.0106	-0.0016	0.00008917
3	-0.75	0	0	0	-0.7853	-0.0037	0.0034	-0.0191	0.00081808
4	0	-0.75	0	0	0.0084	-0.7851	0.0265	-0.0295	0.00143754
5	0.75	0.75	0	0	0.7572	0.7529	0.0077	-0.0104	0.00011385
6	-0.75	0.75	0	0	-0.7806	0.7608	0.0046	-0.0118	0.00060670
7	0.75	-0.75	0	0	0.7745	-0.7818	0.0108	-0.0334	0.00142185
8	-0.75	-0.75	0	0	-0.7888	-0.7964	0.0114	-0.0363	0.00255303
9	0	0	0.75	0	-0.0046	0.0039	0.7853	-0.0115	0.00070736
10	0	0	-0.75	0	-0.0084	-0.0125	-0.7810	-0.0071	0.00061911
11	0	0	0	0.75	-0.0074	-0.0017	0.0114	0.7668	0.00023493
12	0	0	0	-0.75	0.0000	-0.0016	0.0174	-0.7857	0.00078991
13	0.25	0	0	0	0.2875	0.0005	0.0156	-0.0014	0.00082591
14	0	0.25	0	0	0.0016	0.2868	0.0123	-0.0008	0.00075437
15	-0.25	0	0	0	-0.2884	-0.0023	0.0114	-0.0085	0.00084103
16	0	-0.25	0	0	-0.0008	-0.2894	0.0194	-0.0106	0.00102086
17	0.25	0.25	0	0	0.2892	0.2875	0.0130	0.0010	0.00155645
18	-0.25	0.25	0	0	-0.2867	0.2865	0.0085	-0.0048	0.00138722
19	0.25	-0.25	0	0	0.2875	-0.2874	0.0193	-0.0077	0.00161840
20	-0.25	-0.25	0	0	-0.2891	-0.2915	0.0158	-0.0155	0.00187048
21	0	0	0.25	0	-0.0013	0.0013	0.3169	-0.0054	0.00225408
22	0	0	-0.25	0	-0.0007	-0.0038	-0.2904	-0.0036	0.00083003
23	0	0	0	0.25	-0.0009	-0.0012	0.0138	0.2947	0.00109539
24	0	0	0	-0.25	-0.0009	-0.0008	0.0166	-0.3041	0.00160191
25	0.5		0.5		0.5261	-0.0005	0.5519	0.0144	0.00179122
26		0.5	0.5		-0.0011	0.5508	0.5695	-0.0200	0.00390605
27	0.5	0.5	0.5		0.5279	0.5450	0.5414	-0.0053	0.00227273
28	0.5			0.5	0.5230	0.0076	0.0291	0.5452	0.00173831
29		0.5		0.5	0.0012	0.5188	-0.0029	0.5363	0.00084049
30	0.5	0.5		0.5	0.5247	0.5195	0.0115	0.5204	0.00076938
31	0.33333		0.33333	0.33333	0.3668	-0.0053	0.3881	0.3943	0.00393222
32		0.33333	0.33333	0.33333	0.0083	0.3697	0.3876	0.3762	0.00308692
33	0.33333	0.33333	0.33333	0.33333	0.3696	0.3634	0.3779	0.3771	0.00306049
34	1	0.5			0.9119	0.5442	0.0043	-0.0108	0.00492519
35	0.5	1			0.5507	0.9094	0.0110	-0.0096	0.00549601

maximum 0.00549601
average 0.00162718

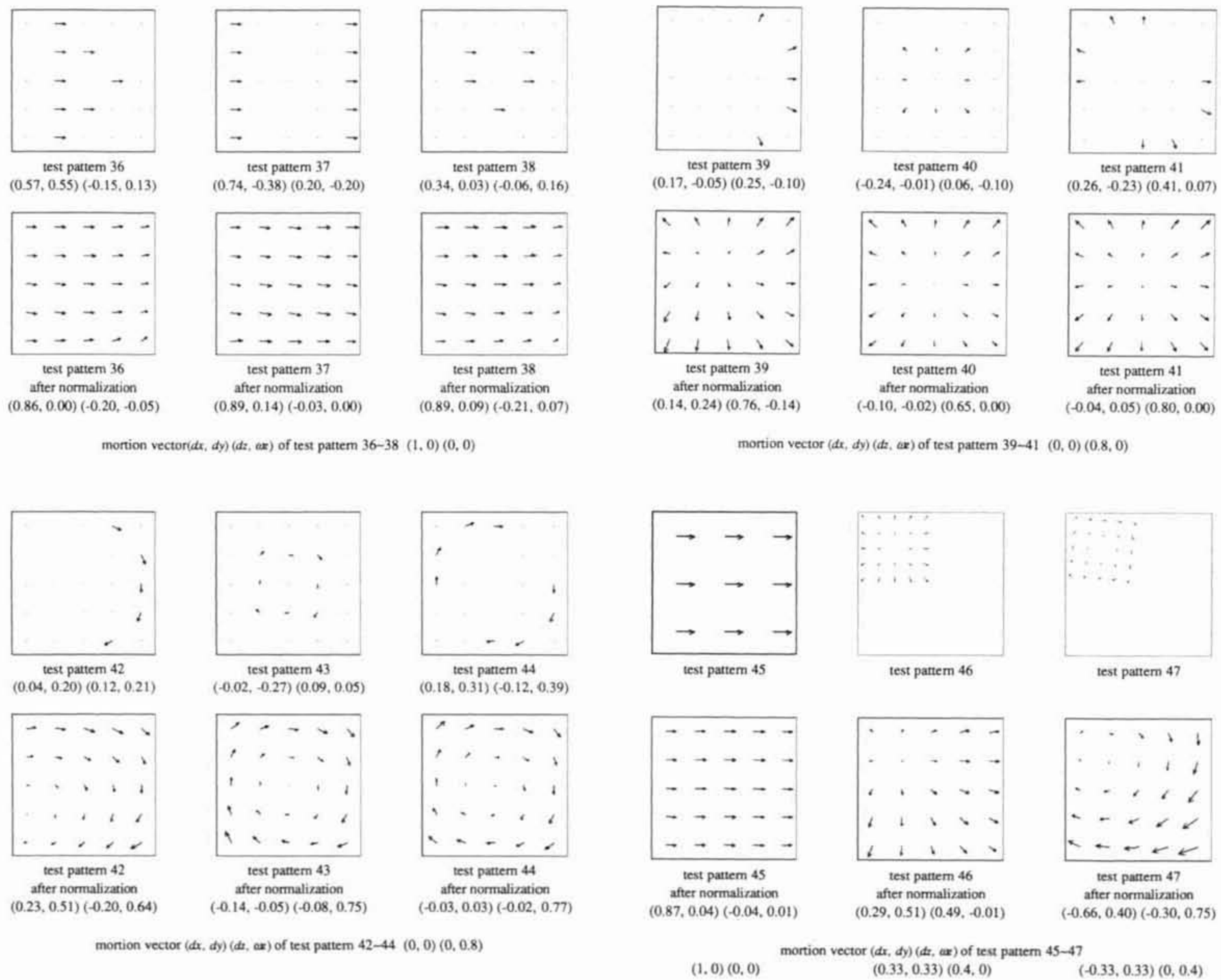


Figure 5 Experimental results on normalization capacity

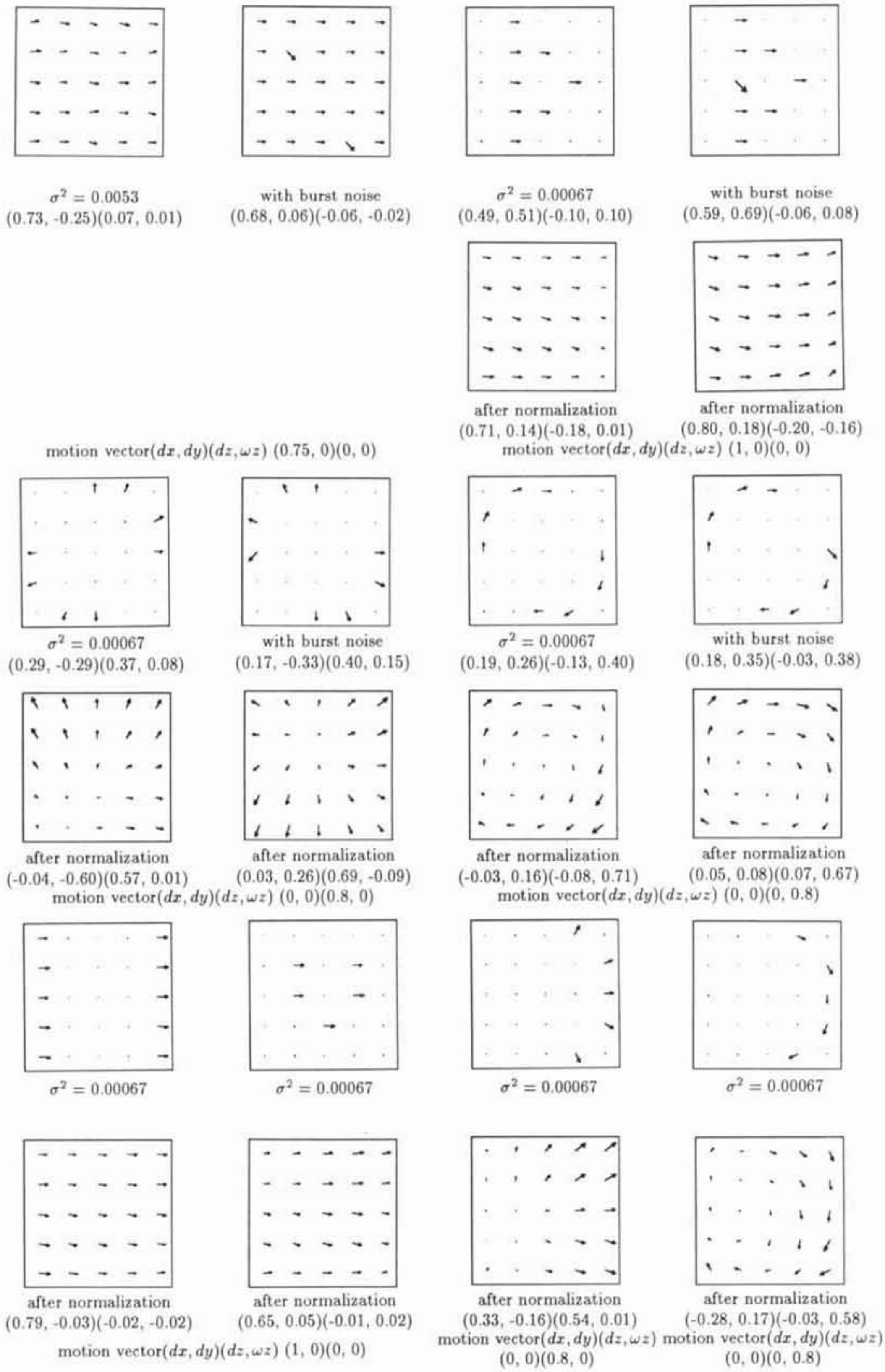
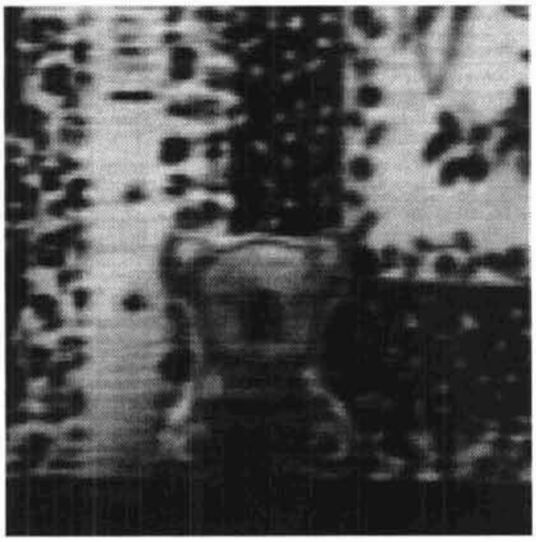
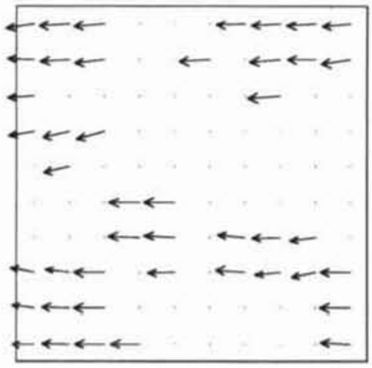


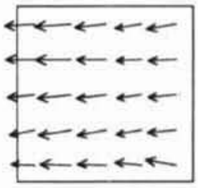
Figure 6 Experimental results with noise added inputs





(a) (b)

Figure 7 Real image and optical flow caluculated from real image sequence



(-2.44, -0.15)(-0.07, -0.20)
motion vector$(dx, dy)(dz, \omega z) \approx (-2.5, 0)(0, 0)$

Figure 8 Experimental results with real image sequence