COMPLEX LOGARITHMIC MAPPING AND THE FOCUS OF EXPANSION[1]

Ramesh Jain

Department of Electrical and Computer Engineering

The University of Michigan

Ann Arbor MI 48109

November 1983

CENTER FOR ROBOTICS AND INTEGRATED MANUFACTURING

Robot Systems Division

COLLEGE OF ENGINEERING

THE UNIVERSITY OF MICHIGAN

ANN ARBOR, MICHIGAN 48109

---

[1]This work was supported by the National Science Foundation, under NSF-G-MCS-8219739.

TABLE OF CONTENTS

Abstract

Complex logarithmic mapping has been shown to be useful for the size, rotation, and projection invariance of objects in a visual field for an observer translating in the direction of its gaze. Assuming known translational motion of the observer, the ego-motion polar transform was successfully used in segmentation of dynamic scenes. By combining the two transforms one can exploit features of both transforms and remove some of the limitations which restrict the applicability of both. In this paper we show that by using complex logarithmic mapping with respect to the focus of expansion rather than the center of the visual field perfect projection invariance and better size and rotation invariance may be obtained for any arbitrary motion of the observer.

1. INTRODUCTION

Recently many researchers [BGT82, CaP77, CAV78, SaT80, SAW72, SCH77, SCH80, SCH81, SCH82] have suggested application of complex logarithmic mapping for the size and rotation invariance of objects in a visual field. Schwartz [SCH77, SCH80, SCH81, SCH82] suggested that the retino-striate mapping can be approximated by complex logarithmic mapping for most visual systems and is responsible for size, rotational, and projection invariance. Cavanagh [CAV78] suggests that a combination of Fourier transform and complex logarithmic mapping may be used for translation, size, and rotational invariance. He argues [CAV81] that the mapping alone results in the invariance only in very limited cases. Many approaches have been proposed for obtaining size and rotation invariance in computer vision systems using either transformations based on complex logarithmic mapping [CaP77, SaT80, SAW72, SWC81, ChW79] or using sensors for acquiring images which have complex logarithmic property.

For the segmentation of dynamic scenes in stationary and nonstationary components of the scene obtained using a moving observer, Jain [JAI82, JAI83] suggested the use of an Ego-Motion Polar (EMP) transform. This transformation uses the known location of the Focus Of Expansion (FOE) as the origin of the polar coordinate system and converts the original image into another rectangular image whose abscissa and ordinates are r and θ, respectively. He developed an algorithm for the segmentation using this transform. The experience with this algorithm [JAI83] shows that if the observer motion is known then the scene can be successfully segmented using the EMP transforms.

The EMP transform has many similarities with the logarithmic mapping. It appears that by combining these two concepts we may be able to understand the scope of the CLM in vision systems better, particularly, for projection invariance. Another advantage may be the possibility of extracting more information from a scene sequence describing the motion of the observer.

In this paper, first, we briefly review the CLM and the EMP and then suggest a modified EMP transformation. We discuss some mathematical aspects of this transform and present results of our experiments to study the efficacy of this transform.

2. COMPLEX LOGARITHMIC MAPPING

The application of complex logarithmic mapping (CLM) has been suggested for implementing a rotation and size invariant mechanism in pattern recognition [CaP77, SAW72]. Schwartz [SCH77, SCH80, SCH81, SCH82] has shown that the mapping from retina to the striate cortex can be approximated by CLM and is responsible for size and rotational invariance in human visual system. Cavanagh [CAV78] suggested that a composite of spatial frequency mapping and CLM would provide a translation, rotation, and size invariant mechanism for human vision. Cavanagh's suggestion was certainly influenced by the work of Casasent and Psaltis [CaP77] in pattern recognition. Schwartz challenged the validity of the Cavanagh's suggestion arguing that application of the Fourier transform leads to certain problems in human visual systems and can not be justified using anatomical arguments also. He posited [SCH81] that the CLM leads to rotation and size invariance and is the mechanism found in most retino-striate mappings. He further showed that in addition this mapping also results in projection invariance, by which he meant the sequence of projective changes that a stimulus fixed in the environment would undergo as an organism approached a fixation point. In a series of papers [SCH77, SCH80, SCH82] Schwartz has shown the importance of the CLM in many tasks of the human visual system.

As argued by Cavanagh [CAV81], the CLM results in invariances only under severe restrictions. The size and rotation invariance is obtained only for those cases when the measurements are made with respect to the origin of the CLM. For the human visual system the CLM represents the retino-striate mapping only in peripheral areas and CLM leads to problems in case of a shift in position of objects. Moreover, the projection invariance claimed by Schwartz is true only when the direction of motion and the axis of gaze are colinear.

In this paper our aim is not to address the relevance of the CLM to the human visual system; but to show that if the FOE is known then the limitation of the CLM in case of projection invariance can be removed and the modified mapping may be used for the extraction of information in the dynamic scenes acquired using a controlled moving camera. Here we discuss some mathematical properties of the transform and in a later section show that by modifying the CLM only slightly it can be made very useful, at least, in the case of a moving observer. The relevance of the proposed modification to the human vision remains to be studied.

2.1. Mathematical Aspects of CLM

Let us consider a point P in an image. If we consider the center of the image as the origin then the cartesian coordinates (x,y) and the polar coordinates (r, θ) of the point are related by:

$$r=\sqrt{(x^2+y^2)} \tag{1}$$

$$\theta=\tan^{-1}(y/x) \tag{2}$$

If it is desired to map a circular region onto a rectangular region, as in case of the human visual system, then it is well known from the theory of complex variables and conformal mapping that an appropriate mapping is:

$$w=log(z) \tag{3}$$

where

$$z=x+iy \tag{4}$$

and

$$w = u(z) + iv(z) \tag{5}$$

are the complex variables. The mapping given by equation 3 simplifies to:

$$u(r,\theta) = \log(r) \tag{6}$$

$$v(r,\theta) = \theta \tag{7}$$

The selection of the above mapping for modelling the retino-striate structure was strongly influenced by the fact that the magnitude of the cortical magnification factor is approximately proportional to retinal eccentricity. It can be easily verified [SCH80] that the proposed mapping satisfies the above relationship as

$$\frac{dw}{dz} = \frac{d(log(z))}{dz} = \frac{1}{r} \tag{8}$$

Another attractive feature of the above mapping is the fact that CLM is the only analytic function which maps an annular region to a rectangular region. The analytic nature guarantees that the mapping preserves the direction and magnitude of the local angles. These properties may be very useful in the extraction of information from the transformed space.

As shown by several researchers [CaP77, SCH80], the rotation of of surfaces in the z space becomes vertical displacement in the w space; the magnification of a surface in z space becomes a horizontal shift in the w space. From dynamic scene analysis view point, a very interesting property of this mapping is that if the observer is moving towards its fixation point then the projection of stationary surfaces remain of the same size and show only horizontal displacements.

3. EGO-MOTION POLAR TRANSFORMATION

The intersection of the 3-D vector representing the instantaneous direction of the observer motion, and the projection plane is called the Focus Of Expansion (FOE). The retinal velocity vectors due to stationary points meet at the FOE. It has been shown [CLO80, GIB79, LEE80, PRA80] that the FOE plays a very important role in the extraction of information from the optical flow.

If the camera motion is known then the FOE can be computed and the fact that the flow vectors for the stationary points meet at the FOE can be used to classify points in an image into stationary and moving points. After computing the optical flow, a test for colinearity of the FOE and the flow vector for the point may be used for the classification. The difficult part in this approach is the computation of the optical flow and applying colinearity test to every point in the image.

If the observer continues its motion in the same direction then the FOE the remains same and the image of a point moves along the radial line originating at the FOE [GIB79, BaB82]. If we transform the original image into another two dimensional rectangular image such that θ is along the conventional Y axis and r is along the conventional X axis, then due to the observer motion the image of a stationary point will display motion only along r in the transformed image. Note that the motion of all stationary points in the original image is in assorted directions but in the transformed image all points show motion in one direction. The surface coherence property and the rigidity assumption allows us to assume that regions in an image representing a surface will show the same motion in the r - θ plane for the translational motion of the observer and the surface. Thus, if the observer has only a translational component to his motion then we can classify all the regions that show only horizontal velocity in the EMP space as due to stationary surfaces. The regions having vertical component of velocity are due to nonstationary surfaces.

Note that the segmentation of a dynamic scene into its stationary and nonstationary components can now be performed by detecting the presence of vertical component of motion in the

EMP space. There is no need to compute optical flow and determine the colinearity of the vectors with the FOE for the segmentation. The velocity components for a region may be determined using techniques used for the stationary camera case in the EMP space. In [JAI82 JAI83] the efficiency of this approach was demonstrated considering several scenes containing stationary and moving objects. In Fig.1 we show three frames from a sequence used in [JAI82b]. The first frame of these frames is shown in EMP form in the Fig.2 and the segmentation obtained using the proposed approach is shown in Fig.3.

Figure 1 Three frames of a sequence acquired using a moving camera. The connecting rod was moving to the left, the yoke to the right and all other objects in the scene were stationary. The camera was moving on a rail.

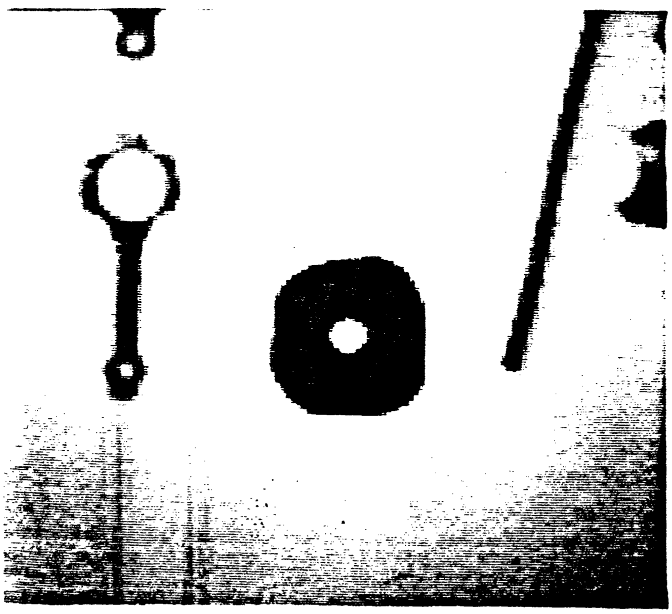
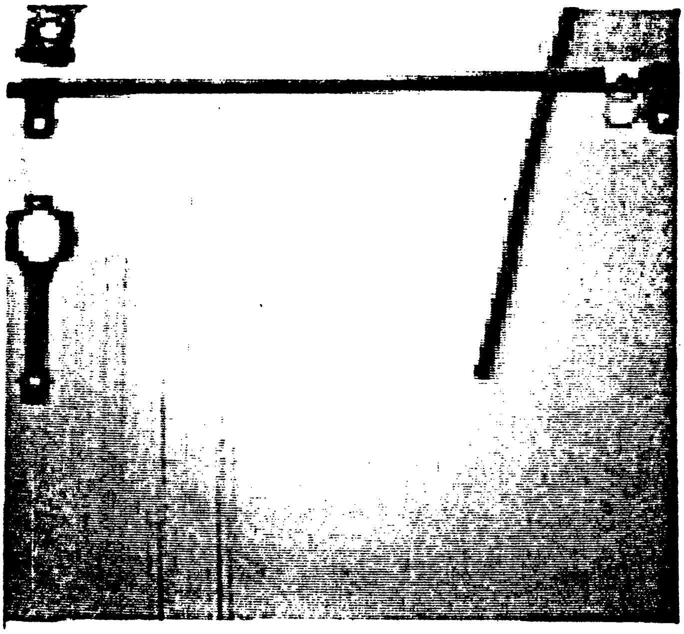


Figure 1 b,c



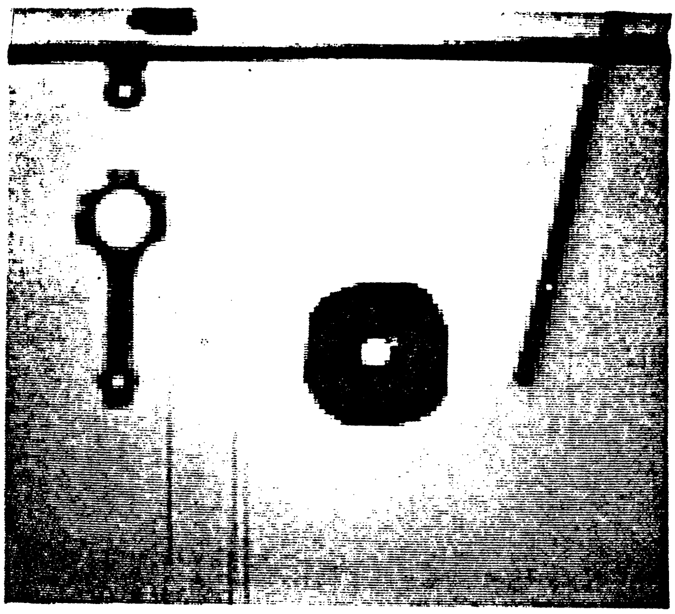


Figure 2 The EMP picture of the frame shown in Figure 1a. Note the distortion in the objects. It should be mentioned here that several issues related to the implementation of the transform are not well understood yet.

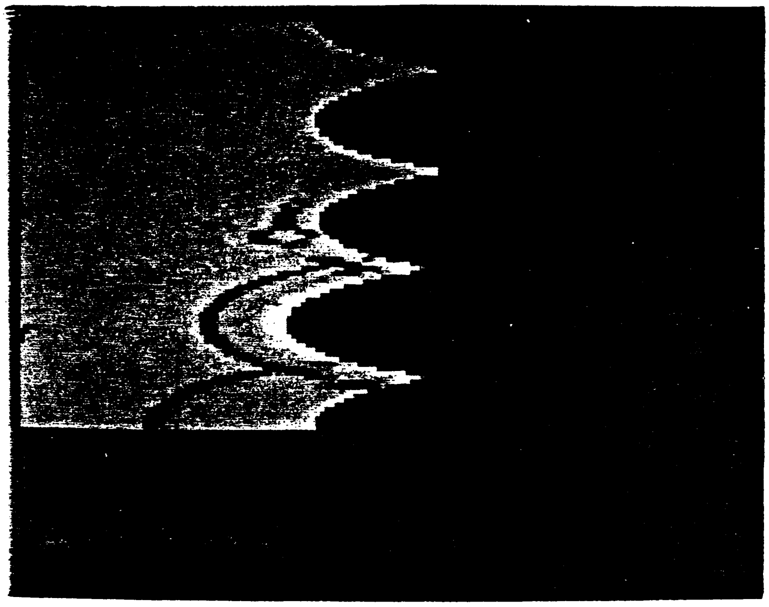
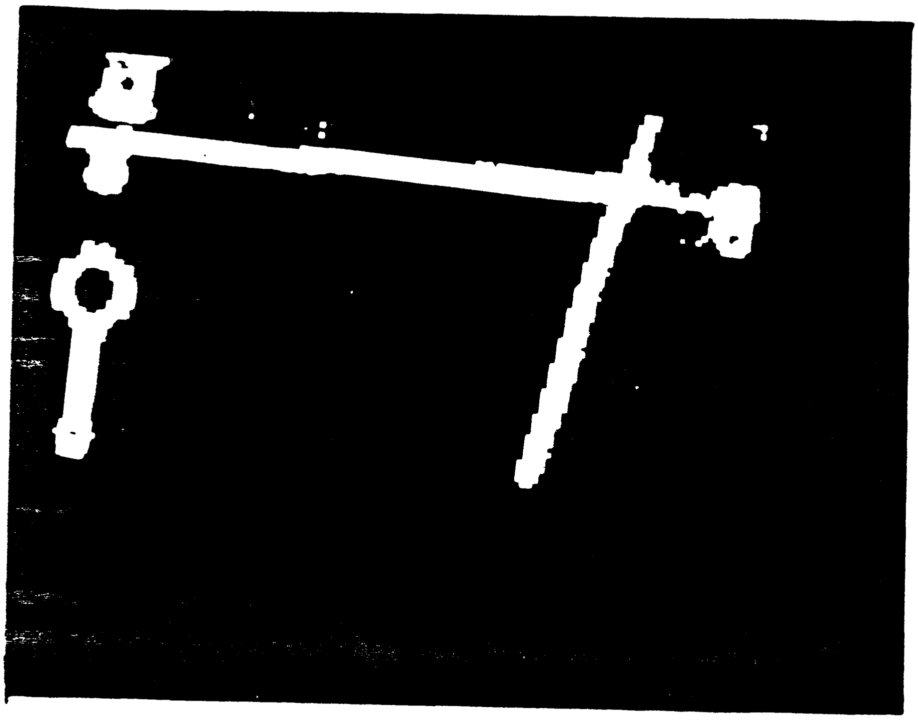


Figure 3 The stationary and moving components of the frame shown in Figure 1a. The brighter objects are the one classified as moving objects and the darker as the stationary object.



4. MODIFICATIONS

Let us consider a stationary point P in the environment whose real world coordinates with respect to the observer at a time instant are (X,Y,Z). The projection of this point on the image plane, assuming that the projection plan is parallel to XY plan and is at $Z=1$, is at (x',y') given by

$$x' = \frac{X}{Z} \tag{9}$$

$$y' = \frac{Y}{Z} \tag{10}$$

If the observer moves then the relationship between the image plane distance of the projection of the point from the FOE and the distance Z of the point from the observer is

$$\frac{dr}{dZ} = \frac{d(\sqrt{(x'^2+y'^2)})}{dZ} = -\frac{r}{Z} \tag{11}$$

where r is the distance from the center of the image.

The projection invariance shown by Schwartz is based on the fact that for a moving observer the projection of a point at the distance Z from the observer satisfies the following relations in the W plane:

$$\left|\frac{dr}{dZ}\right| = \left|\frac{r}{Z}\right| \tag{12}$$

The complex logarithmic mapping results in

$$\frac{du}{dZ} = \frac{du}{dr} * \frac{dr}{dZ} \tag{13}$$

and

$$\frac{dv}{dZ} = \frac{dv}{d\theta} * \frac{d\theta}{dZ} \tag{14}$$

since from equation 6, we get

$$\frac{du}{dr} = \frac{1}{r} \tag{15}$$

and

$$\frac{d\theta}{dZ} = \frac{d(\tan^{-1}\frac{y'}{x'})}{dZ} = 0 \tag{16}$$

in CLM; we have

$$\frac{du}{dZ} = -\frac{1}{Z} \tag{17}$$

$$\frac{dv}{dZ} = 0 \tag{18}$$

Thus for a stationary point P the displacement along v is zero for a translating observer in the direction of its fixation point. Moreover, if the stationary point is far away from the observer its displacement in the CLM image may be considered uniform for the uniform translatory motion of the observer. If the observer moves in the direction of its gaze, then the displacement component (dx, dy, dz) is such that $dx = dy = 0$ resulting in the FOE at the center of the visual field. As argued by Cavanagh [CAV81], if the direction of the motion of the observer is not in the direction of its gaze then the above relationship will be no longer correct. In this case the FOE will be at the point $(dx/dz, dy/dz)$ and the above equations are satisfied if the CLM is obtained considering the FOE as the origin for the transformation.

The EMP transformation considers the direction of motion of the observer by using the FOE as the origin but does not use CLM. By using CLM with respect to the FOE rather than the center of the image we can combine features of both approaches.

Another important point to be noted is that the displacement of the projection of the point in the transformed space is inversely proportional to the depth of the point. The rate of change of the displacement is thus

$$\frac{d^2u}{dZ^2} = \frac{1}{Z^2} \tag{19}$$

Clearly, the displacement and all its derivatives of the projection of a point in W space are governed by the 3-D distance of the point in the real space. Thus for a planar surface perpendicular to the direction of the gaze the displacement of every point in the transformed space will be same; for other surface the displacements of different points will depend on the nature and the orientation of the surface. This fact may be exploited for the determination of the orientation of a planar surface. It appears possible to obtain the nature of the non-planar surfaces by determining the displacement component at every point of the surface and then using techniques similar to those used in the photometric stereo.

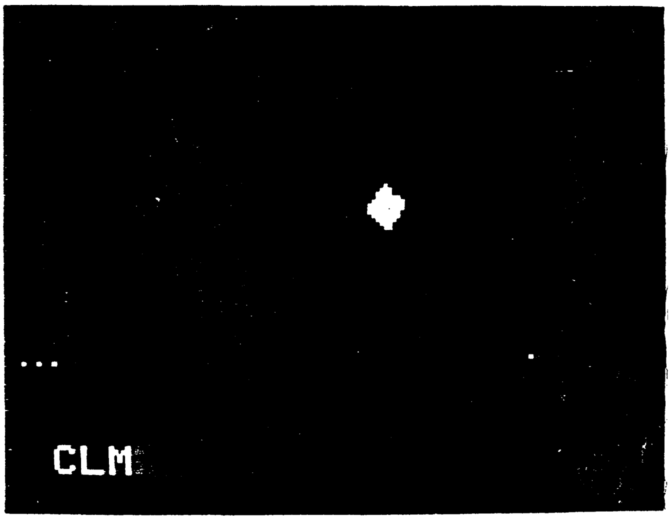
5. RESULTS

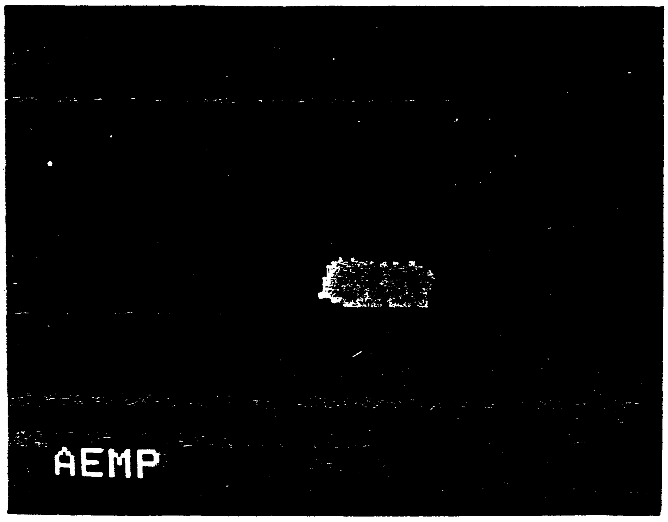
We simulated motion of points at different distances using the perspective transformation. The CLM with respect to the center of the image and with respect to the FOE for the uniform observer motion for several frames are given in Table 1. It can be clearly seen that the angle remains constant when the CLM is taken with respect to the FOE; but varies in the other case. Our experiments on IBM PC showed that in CLM the departure in the angle depends on the relative values of X,Y,Z and dx,dy,dz. From the table we see that the departure was more for points having lower values of X and Z; since dy=0, the Y component has no effect. In the modified transform, labelled EMP in the table, the angle remained constant.

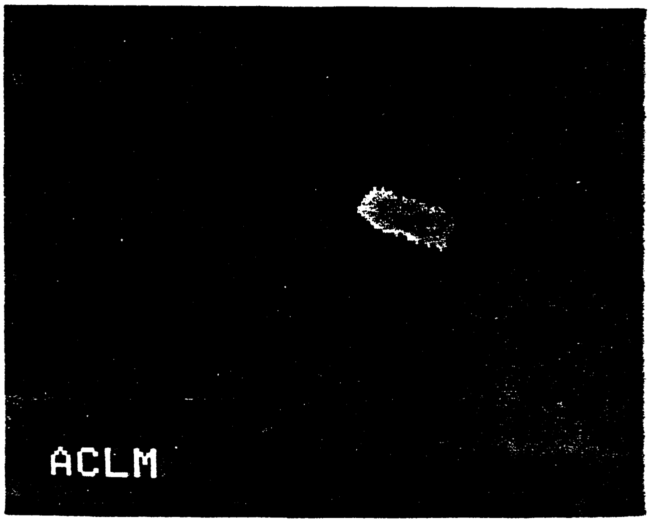
Another interesting observation to be made from this table is the fact that points at the same depth, i.e., equal Z, show equal amount of displacement in the EMP space; not in the CLM space. The DCR and DER columns in the table show the frame-to-frame displacement of points in the CLM and EMP spaces, respectively. Thus, if it is desired to extract the surface information using the known motion of the observer then EMP space retains the information; the CLM introduces noise. Of course in depth extraction using motion stereo one can control camera motion, such that the camera moves in the direction of its optical axis resulting in the FOE at the center of the image, then CLM and EMP will be same. In more general cases, however, use of the observer motion information is desirable.

In Figure 4 we show motion of the observer with respect to a simulated surface. The FOE is (20,0) from the center of the image of size 128x128. The figure shows the surface in the CLM and the modified EMP space and also the superimposed images of the surface after several frames. Note that image plane motion of the surface in the modified EMP space is horizontal, in the CLM space is not. This figure demonstrates that the projection invariance is obtained in the EMP space, not in the CLM space.

Figure 4 The motion of a surface in the CLM and EMP spaces due to the motion of the observer. The FOE is (20,0) from center in the image of the size 128x128.







6. DISCUSSION

In this paper we showed that by using the CLM with respect to the FOE the applicability of the CLM in dynamic scenes can be significantly increased. We considered a very simple simulation to illustrate the enhancements in the projection invariance using the modified CLM (or modified EMP). In more realistic situations involving complex motion of observer and planar and curved surfaces in the scene we intend to compute *some* properties of neighborhood of points to extract information about the nature of surfaces.

The plausibility of the modified transformation in the human visual system has not been investigated here. We are investigating the possibility of a simple transform based on the ego-motion to be applied to the CLM to obtain the modified transform. If such a transform can be obtained then the presence of such a transform from the striate cortex to the higher levels should be investigated.

REFERENCES

[BaB82]

Ballard, D.H. and C. M. Brown, *Computer Vision,* Prentice Hall, 1982.

[BGT79]

Braccini, C., G. Gamberdella, and V. Tagliasco, "A model of the early stages of human visual system," *Biological Cybernetics,* 44, 1982, pp.47-88

[CaP77]

"New optical transforms for pattern recognition," Casasent, D. and D. Psaltis, *Proc. of IEEE,* vol. 65, pp.77-84, 1977.

[Cav78]

Cavanagh, P., "Size and position invariance in the visual system," *Perception,* vol. 7, pp.167-177, 1978.

[Cav81]

Cavanagh, P., "Size invariance: reply to Schwartz," *Perception,* col. 10, pp.469-474, 1981.

[ChW79]

Chaikin, G. and C. Weiman, "Log spiral grids in computer pattern recognition", *Computer Graphics and Pattern Recognition,* vol.4, pp.197-226, 1979.

[Clo80]

Clocksin, W.F., "Perception of surface slant and edge labels from optical flow: A computational approach," *Perception,* vol. 9, 1980, pp.253-269.

[Gib79]

Gibson, J.J., *The ecological approach to visual perception,* Houghton Mifflen, Boston, 1979.

[Jai82]

Jain, R., "Segmentation of moving observer frame sequences," *Pattern Recognition Letters,* vol. 1 , pp. 115-120, 1982.

[Jai83]

Jain, R. "Segmentation of frame sequences obtained by a moving observer," *General Motors Research Publication,* no. 4247, Jan. 1983.

[Law82]

Lawton, D. T., "Motion analysis via local translational processes," *Proc. Computer Vision Workshop, 1982, pp. 59-72.*

[Lee80]

Lee, D.N., "The optic flow field: The foundation of vision," *Phil. Trans. Royal Society of London,* vol. B290, 1980, pp. 169-179.

[Pra80]

Prazdny, K., "Egomotion and relative depth map from optical flow," *Biological Cybernetics,* vol. 36, 1980, pp. 87-102.

[SaT80]

Sandini, G and V. Tagliasco, "An anthropomorphic retin-like structure for scene analysis," *Computer Graphics and Image Processing,* vol.14, pp.365-372, 1980.

[SAW72]

Sawchuk, A. A., "Space-variant image motion degradation and restoration", *Proc. IEEE,* vol. 60, pp.854-861, 1972.

[SCH77]

Schwartz, E. L., "The development of specific visual connections in the monkey and goldfish: Outline of a geometric theory of receptotopic structure", *J. Theoretical Biology,* vol 69, pp.655-683.

[Sch80]

Schwartz, E. L., "Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to coding," *Vision Research,* 20, 1980, pp.645-669.

[Sch81]

Schwartz, E. L., "Cortical anatomy, size invariance, and spatial frequency analysis," *Perception,* vol.10, pp.455-468, 1981.

[Sch82]

Schwartz, E.L., "Columnar architecture and computational anatomy in primate visual cortex: Segmentation and feature extraction via spatial frequency coded difference mapping," *Biological Cybernetics,* vol. 42, pp.157-168, 1982.

[SWC81]

Schenker, P.S., K.M. Wong, and E.G.Cande "Fast adaptive algorithms for low-level scene analysis: Application of polar exponential grid (PEG) representation to high-speed, scale-and-rotation invariant target segmentation", *Proc. SPIE,* Vol. 281, Techniques and Applications of Image Understanding, pp.47-57, 1981.

Table 17

Table 1

FOR THIS RUN NO OF POINTS WAS 4

THE 3-D COORDINATES OF POINTS WERE

1	150	300	15
2	750	300	15
3	2000	3000	50
4	400	700	50

OBSERVER MOTION 3 0 1

THE FOCUS OF EXPANSION IS 3 0

FOR POINT NO. 1

FRAME	CLM_R	CLM_T	DCR	DCT	EMP_R	EMP_T	DER	DET
	--FOR COMPLEX LOG MAPPING			----	---FOR EGO-MOTION LOG POLAR--			
1.00000	3.17232	1.11518	0.00000	0.00000	3.12250	1.23412	0.00000	0.00000
2.00000	3.24250	1.12328	0.07018	0.00810	3.19661	1.23412	0.07411	0.00000
3.00000	3.31867	1.13144	0.07617	0.00816	3.27666	1.23412	0.08004	0.00000
4.00000	3.40186	1.13966	0.08319	0.00822	3.36367	1.23412	0.08701	0.00000
5.00000	3.49340	1.14794	0.09154	0.00828	3.45898	1.23412	0.09531	0.00000
6.00000	3.59504	1.15629	0.10165	0.00835	3.56434	1.23412	0.10536	0.00000
7.00000	3.70917	1.16470	0.11413	0.00841	3.68212	1.23412	0.11778	0.00000
8.00000	3.83910	1.17317	0.12993	0.00847	3.81565	1.23412	0.13353	0.00000
9.00000	3.98971	1.18170	0.15061	0.00853	3.96980	1.23412	0.15415	0.00000
10.00000	4.16855	1.19029	0.17884	0.00859	4.15212	1.23412	0.18232	0.00000

FOR POINT NO. 2

FRAME	CLM_R	CLM_T	DCR	DCT	EMP_R	EMP_T	DER	DET
	--FOR COMPLEX LOG MAPPING			----	---FOR EGO-MOTION LOG POLAR--			
1.00000	4.05177	0.38189	0.00000	0.00000	4.00235	0.40232	0.00000	0.00000
2.00000	4.12242	0.38328	0.07064	0.00139	4.07646	0.40232	0.07411	0.00000
3.00000	4.19899	0.38469	0.07657	0.00140	4.15650	0.40232	0.08004	0.00000
4.00000	4.28251	0.38610	0.08353	0.00141	4.24352	0.40232	0.08701	0.00000
5.00000	4.37433	0.38752	0.09182	0.00142	4.33883	0.40232	0.09531	0.00000
6.00000	4.47619	0.38896	0.10186	0.00143	4.44419	0.40232	0.10536	0.00000
7.00000	4.59046	0.39040	0.11427	0.00144	4.56197	0.40232	0.11778	0.00000
8.00000	4.72046	0.39185	0.13001	0.00145	4.69550	0.40232	0.13353	0.00000
9.00000	4.87108	0.39332	0.15062	0.00146	4.84965	0.40232	0.15415	0.00000
10.00000	5.04986	0.39479	0.17878	0.00147	5.03197	0.40232	0.18232	0.00000

FOR POINT NO. 3

FRAME	CLM_R	CLM_T	DCR	DCT	EMP_R	EMP_T	DER	DET
	--FOR COMPLEX LOG MAPPING ----				---FOR EGO-MOTION LOG POLAR--			
1.00000	4.29795	0.98349	0.00000	0.00000	4.27569	1.01821	0.00000	0.00000
2.00000	4.31811	0.98418	0.02016	0.00069	4.29631	1.01821	0.02062	0.00000
3.00000	4.33870	0.98487	0.02059	0.00069	4.31736	1.01821	0.02105	0.00000
4.00000	4.35974	0.98557	0.02105	0.00069	4.33887	1.01821	0.02151	0.00000
5.00000	4.38126	0.98626	0.02152	0.00070	4.36085	1.01821	0.02198	0.00000
6.00000	4.40328	0.98696	0.02201	0.00070	4.38332	1.01821	0.02247	-0.00000
7.00000	4.42581	0.98766	0.02253	0.00070	4.40631	1.01821	0.02299	0.00000
8.00000	4.44888	0.98835	0.02307	0.00070	4.42984	1.01821	0.02353	0.00000
9.00000	4.47252	0.98905	0.02364	0.00070	4.45394	1.01821	0.02410	0.00000
10.00000	4.49675	0.98975	0.02423	0.00070	4.47863	1.01821	0.02469	0.00000

FOR POINT NO. 4

FRAME	CLM_R	CLM_T	DCR	DCT	EMP_R	EMP_T	DER	DET
	--FOR COMPLEX LOG MAPPING ----				---FOR EGO-MOTION LOG POLAR--			
1.00000	2.79870	1.05489	0.00000	0.00000	2.71928	1.22777	0.00000	0.00000
2.00000	2.81748	1.05814	0.01878	0.00325	2.73990	1.22777	0.02062	0.00000
3.00000	2.83671	1.06140	0.01922	0.00326	2.76096	1.22777	0.02105	0.00000
4.00000	2.85640	1.06467	0.01969	0.00327	2.78246	1.22777	0.02151	0.00000
5.00000	2.87656	1.06795	0.02017	0.00328	2.80444	1.22777	0.02198	0.00000
6.00000	2.89723	1.07125	0.02067	0.00330	2.82691	1.22777	0.02247	0.00000
7.00000	2.91842	1.07456	0.02119	0.00331	2.84990	1.22777	0.02299	0.00000
8.00000	2.94016	1.07788	0.02174	0.00332	2.87343	1.22777	0.02353	-0.00000
9.00000	2.96247	1.08121	0.02231	0.00333	2.89753	1.22777	0.02410	0.00000
10.00000	2.98539	1.08455	0.02292	0.00334	2.92223	1.22777	0.02469	0.00000