Global versus Structured Interpretation of Motion: Moving Light Displays

Jeffrey E. Boyd
Department of Electrical and Computer Engineering
University of California, San Diego, La Jolla CA 92093-0407
jeffboyd@ece.ucsd.edu

James J. Little
Department of Computer Science, University of British Columbia
Vancouver, B.C., Canada V6T 1Z4
little@cs.ubc.ca

January 22, 1997

Abstract

Moving light displays (MLDs) have been used extensively to study motion perception and perception of the human gait in particular. MLD perception is largely considered to be structural, i.e., perception depends on identification of human kinematic structure. However, recent work by Little and Boyd has shown that it is possible to recognize individual people, from their gaits, by non-structural means. They use global *shape-of-motion* features derived from optical flow in a sequence of gray-scale images. Our goal is to show that shape-of-motion features can be derived equally well from MLD images as from gray-scale images, and to compare the recent results obtained for shape-of-motion recognition with psychophysical observations about MLD perception. The implication is that non-structural shape-of-motion interpretation of gait can be applied to MLDs, allowing us to interpret significant MLD results in the context of a known algorithm. Our results shed light on the validity of shape-of-motion features from the psychophysical standpoint as well as suggest an alternative approach to understanding MLD perception. In particular, we find that characterizing movement in a gait may be treated as the sum of a set of moving points (if this is true then MLD lights need not be placed right at joints). Changes to a subset of the points affect the sum and consequently affect the perception of the whole.

Keywords: nonrigid/articulated motion estimation, synthesis and understanding, moving light displays, optical flow, gait, perception

Global versus Structured Interpretation of Motion: Moving Light Displays
Summary

(A) What is the original contribution of the paper? This paper contributes an explanation of psychophysical observations regarding moving light displays in the context of non-structural shape-of-motion features. The result is that characterizing motion in a gait may be treated as a process of summation, where the sums describe the distribution of optical flow, and the phase relationships important for recognition are preserved in the sums.

(B) What is the most closely related work by others? The most closely related work in the computer vision literature is that of Polana and Nelson [12, 14]. In the psychophysical literature, the most recent and relevant work is that of Bertenthal and Pinto [4].

(C) Has this paper been submitted elsewhere? No.

Global versus Structured Interpretation of Motion: Moving Light Displays

Abstract

Moving light displays (MLDs) have been used extensively to study motion perception and perception of the human gait in particular. MLD perception is largely considered to be structural, i.e., perception depends on identification of human kinematic structure. However, recent work by Little and Boyd has shown that it is possible to recognize individual people, from their gaits, by non-structural means. They use global *shape-of-motion* features derived from optical flow in a sequence of gray-scale images. Our goal is to show that shape-of-motion features can be derived equally well from MLD images as from gray-scale images, and to compare the recent results obtained for shape-of-motion recognition with psychophysical observations about MLD perception. The implication is that non-structural shape-of-motion interpretation of gait can be applied to MLDs, allowing us to interpret significant MLD results in the context of a known algorithm. Our results shed light on the validity of shape-of-motion features from the psychophysical standpoint as well as suggest an alternative approach to understanding MLD perception. In particular, we find that characterizing movement in a gait may be treated as the sum of a set of moving points (if this is true then MLD lights need not be placed right at joints). Changes to a subset of the points affect the sum and consequently affect the perception of the whole.

Keywords: nonrigid/articulated motion estimation, synthesis and understanding, moving light displays, optical flow, gait, perception

1 Introduction

Moving light displays (MLDs) have been used extensively to study the perception of motion. In a moving light display, lights, or bright reflecting patches, are fastened to points on rigid objects and images of the objects are recorded as they move. The result is an image sequence where the objects themselves are not visible, only the lights. A major use of MLDs has been in the study of gait perception [1, 4, 7, 8, 15]. In this case, the lights are fastened to the joints of a pedestrian. In the resulting image, the pedestrian is invisible, only the joint positions can be seen. In general, psychophysical research has shown that much of the information required to interpret a moving sequence is captured in an MLD.

Throughout the psychophysical and computer vision literature, MLD perception is largely considered to be structurally based, i.e., perception depends on identification of structural components of the pedestrian's body such as shins, forearms, torsos and joints [6]. While this interpretation is not universal, it certainly predominates. Alternative approaches avoid structure, but track points over time and analyze the movement of the points.

Not all analysis of motion depends on MLDs. Polana and Nelson [12, 13, 14], in a non-structural approach, look at global spatial distributions of motion for a figure engaged in some activity. They compute spatial statistics in a coarse mesh and derive a vector describing the relative magnitudes and periodicity of activity in regions, over time. Their experiments demonstrate that the values so derived can be used to discriminate among differing activities.

Baumberg and Hogg [2] present a method of representing the global shape of a moving body at an instant in time. Their method produces a description composed of a set of principal spline components and a direction of motion. In later work [3], Baumberg and Hogg add temporal variation by modeling the changing shape as a vibrating plate. They create a vibration model for a *generic* pedestrian and then measure the quality of fit of the generic data to another pedestrian.

In contrast, a recent example of a structured approach is that of Ju, Black and Yacoob [9]. They model the structure of the human body as pieces of a *cardboard person model.* A person's limbs are modeled as connected planar patches where patches represent body parts such as shins or thighs. They compute optical flow using this model while incorporating structural constraints about joints, e.g., they force the upper part of the shin to be near the lower part of the thigh.

Recent work by Little and Boyd [10, 11], in the spirit of Polana and Nelson, and Baumberg and Hogg, has shown that it is possible to recognize individual people, from their gaits, by non-structural means. The method uses global *shape-of-motion* features derived from optical flow patterns in a sequence of gray-scale images. They measure a set of scalar values that characterize the shape of the entire moving region of a pedestrian in each frame of a motion sequence. These values, taken as a time series, vary cyclically with the gait. Little and Boyd use the phase relationships between these series to discriminate amongst a set of individual subjects of an experiment. Although shape-of-motion features vary between individuals to the extent that recognition is possible, they still only vary over a small range for the entire population. For a pedestrian the features tend to fall into a certain limited range, while individuals vary within that range.

Our goal here is to show that shape-of-motion features can be derived equally well from MLD images as from gray-scale images. In doing so, we compare the recent results obtained for recognition from optical flow with observations about MLD gait recognition reported in the psychophysical literature. The implication is that non-structural, non-tracking, interpretation of gait using global features can be applied to MLDs too. This allows us to interpret significant psychophysical MLD results in the context of a known algorithm. The results throw some light onto the validity of shape-of-motion features from the psychophysical standpoint as well as suggest an alternative approach to understanding psychophysical observations about MLD perception. In particular, we find that characterizing movement in a gait may be treated as the sum of a set of moving points. Changes to a subset of the moving points affect the sum and consequently affect the perception of the whole. If this model is correct, then it is not necessary for the moving lights to be placed at joints to perceive gait, just that they adequately sample the distribution of motion.

The following section reviews some significant psychophysical results for MLD perception. We then go on to give a brief description of shape-of-motion features and show how they are used to characterize and recognize gaits. We describe an experiment that compares shape-of-motion features generated from a gray scale sequence to those derived from the equivalent MLD. The results of the experiment show that most of the shape-of-motion features are about the same for both types of input sequences. We then discuss the implications of this result with respect to observations made in the psychophysical literature.

2 Perception of Moving Light Displays

Perception of motion has long been of interest to researchers. In the psychophysical literature, much of the research has focused on isolating the stimuli that give certain perceptions. A common method is to use MLDs. Cedras and Shah [6] present a survey of motion-based recognition in the field of computer vision. They identify two theories about interpretation of MLDs. The first says that people use motion to recover three-dimensional structure and then recognize using that structure. The other theory suggests that motion information is used directly. In this section we review some significant results in gait perception using MLDs. While most explanations for gait perception are of the structural variety, we go on to show that some recent results of Little and Boyd [10, 11] in non-structural interpretation from gray-scale images may apply to MLDs and suggest new interpretations of psychophysical observations.

Johansson [7, 8] pioneered the use of MLDs in perception. In experiments, he marked the joint positions of a person with lights and recorded the images as the person walked in front of a camera. He also used bright patches on joints instead of lights and recorded the images with high contrast to allow more freedom of movement. Johansson observed that viewers of the generated pattern of dots failed to recognize the static pattern but quickly recognized a pedestrian from a moving sequence. Johansson concluded that the recognition was structural. This appears to be inspired by Johansson's work with simpler patterns of dots where the assumption of rigidity and perspective projection explained perceptions of structure.

Barclay, Cutting and Kozlowski [1] tested for human recognition of gender from MLDs. Their results showed that some gender recognition is possible. They report that with a 4.4 second sequence, 66% of 57 subjects identified the gender of the pedestrian correctly. Recognition deteriorated with shorter sequences. While the results do indicate some gender recognition, it is not much better than chance, 50%, and one should not assume that, in general, gender is recognizable from MLDs.

Sumi [15] conducted experiments that examined perceptions of MLDs for gaits when the MLD is inverted. Interestingly, he concludes that "the majority of the subjects recognized it always as the upright-forward human movement." However, the motion did not appear to be normal to subjects either and "the pattern was perceived mostly as a very strange and comical action of the human body." This suggests that properties of motion required for recognizing a human gait are at least partly preserved when the image is inverted, but certainly not entirely.

Bertenthal and Pinto [4] identify the importance of phase in gait perception. They model the motion of the human body as a hierarchical organization of pendulums where each pendulum is a rigid link in a limb. In this context, they refer to the phases of the oscillations of these links. To investigate the significance of phase, they create MLD images of a canonical pedestrian, masked with a pattern of dots. They then perturb the phase of the oscillations of a limb and test for gait perception for both the canonical and phase-perturbed sequences. The results of their experiment showed that "the canonical target was significantly better than the detection of the phase-perturbed target; however, detection of both targets was significantly above chance." The phase Berthenthal and Pinto refer to is not the phase that Little and Boyd speak of, but the two are related as shown in Section 6.1. Phase, whether it is the phase of a limb's oscillation or the phase of a shape-of-motion feature, is definitely crucial to gait perception and recognition.

Most researchers using MLDs have sought a structural explanation for the perceptions observed. Johansson [8] projected the rigid structural interpretation of a pattern of a few dots to the more complex motion of dots. Bertenthal and Pinto [4] look at phases of structural components such as limbs. In contrast, Little and Boyd have shown that individual pedestrians can be recognized by a machine using non-structural shape-of-motion features. These characteristics are derived from optical flow computed from a sequence of gray-scale images. The remainder of this paper shows that shape-of-motion features can be derived readily from MLDs, and as a consequence, shape-of-motion features can explain much of what has been observed about gait perception, without relying on structure.

3 Shape-of-Motion Features for Recognition

Previous work by Little and Boyd [10, 11] shows that recognition of humans from their gaits is possible using non-structural means. Figure 1 illustrates the data flow through the system that creates the shape-of-motion features that are used for recognition [11]. The system begins with an image sequence of $n+1$ images featuring the frontoparallel motion of a single pedestrian walking in front of a static background, and then derives n dense optical flow images. For each of these optical flow images, the system computes m characteristics that describe the shape of the motion (i.e., the spatial distribution of the flow), for example, the centroid of the moving points, and various moments of the flow distribution. Some of these are locations in the image, but all are treated as time-varying scalar values. Table 1 summarizes the scalar values used. Rearranging the scalar values forms a time series for each scalar. A walking pedestrian undergoes periodic motion, returning to a standard position after a certain time period that depends on the frequency of the gait. The system analyzes the periodic structure of these time series and determines the fundamental frequency of the variation of each scalar. The set of time series for a view shares the same frequency, or simple multiples of the fundamental, but their phases vary. To make different sequences comparable, the system subtracts a reference phase, ϕ_m, derived from one of the scalars. Each image sequence is characterized by a vector, $F = (F_1, \ldots, F_{m-1})$, of $m-1$ relative phase features. The phase feature vectors are then used to recognize individuals.

The process was applied to 25 different gait sequences, five sequences for each of five subjects. Analysis of variance of the phase features indicates that there is significant variation of several of the features between subjects. Assured that there was more than just random variation in the phases, Little and Boyd used the phase features to test recognition. For each subject, they computed an exemplar feature by finding the average of each phase feature for that subject. They then found the best match for a sequence from among the exemplars. Using only three features with significant variations and appropriate cross validation techniques they achieved recognition rates of about 90%.

At no point in shape-of-motion recognition is the structure of the walking subject recovered. The process operates entirely without a model of human kinematics. While the method computes dense optical flow from gray-scale images, optical flow from an equivalent

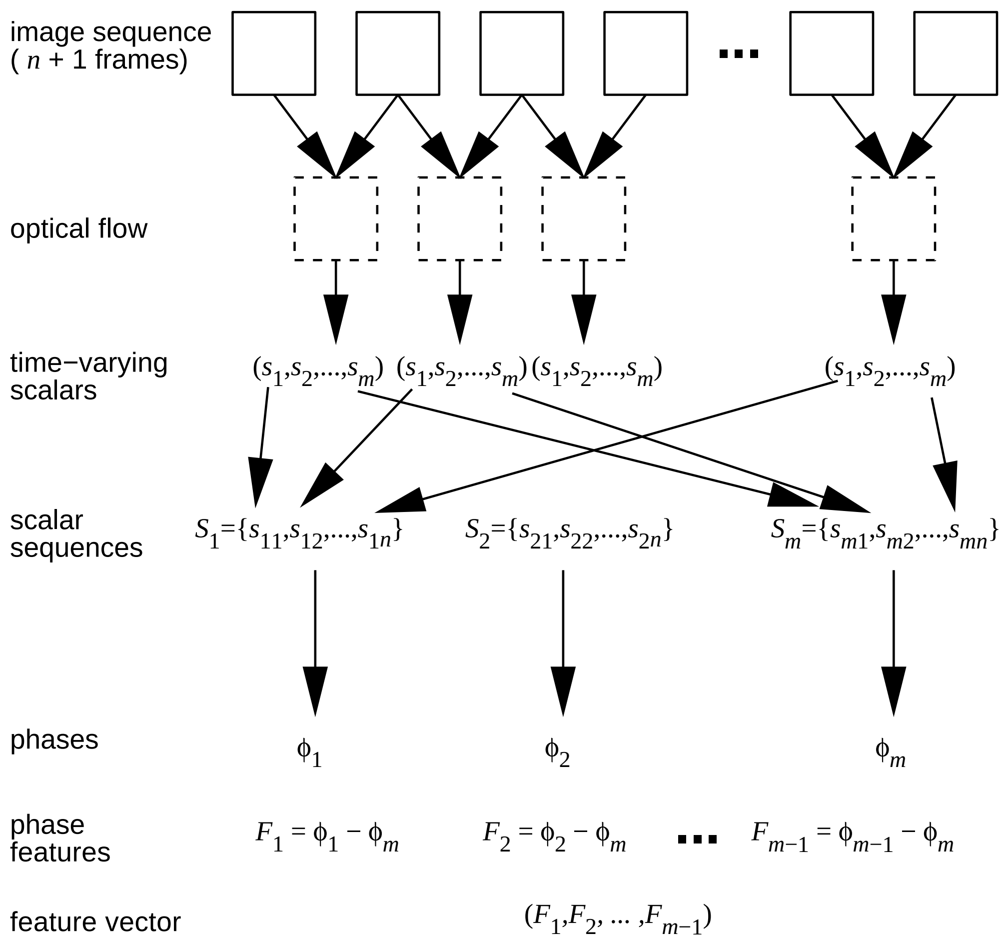


Figure 1: The structure of shape-of-motion image analysis. Each image sequence produces a vector of $m-1$ phase values used as features for recognition.

Description	Label	Formula
x coordinate of centroid,	x_c	$\sum xT / \sum T$
y coordinate of centroid,	y_c	$\sum yT / \sum T$
x coordinate, centroid of $\|(u,v)\|$ distribution	x_{wc}	$\sum x\|(u,v)\|T / \sum \|(u,v)\|T$
y coordinate, centroid of $\|(u,v)\|$ distribution	y_{wc}	$\sum y\|(u,v)\|T / \sum \|(u,v)\|T$
x coordinate of difference of centroids	x_d	$x_{wc} - x_c$
y coordinate of difference of centroids	y_d	$y_{wc} - y_c$
aspect ratio (or elongation) – ratio of length of major axis to minor axis of an ellipse	a_c	$\lambda_{max}/\lambda_{min}$, where λs are eigenvalues of second moment matrix for motion distribution
elongation of weighted ellipse	a_{wc}	as in a_c, but for weighted distribution
difference of elongations,	a_d	$a_c - a_{wc}$
x coordinate, centroid of $\|u\|$ distribution	x_{uwc}	$\sum x\|u\|T / \sum \|u\|T$
y coordinate, centroid of $\|u\|$ distribution	y_{uwc}	$\sum y\|u\|T / \sum \|u\|T$
x coordinate, centroid of $\|v\|$ distribution	x_{vwc}	$\sum x\|v\|T / \sum \|v\|T$
y coordinate, centroid of $\|v\|$ distribution	y_{vwc}	$\sum y\|v\|T / \sum \|v\|T$

Table 1: Summary of scalar shape-of-motion descriptors. Summations are over the entire image. u and v are the x- and y-direction optical flow values respectively. The function T segments the image. $T = 1$ for pixels that are moving and $T = 0$ for stationary pixels.

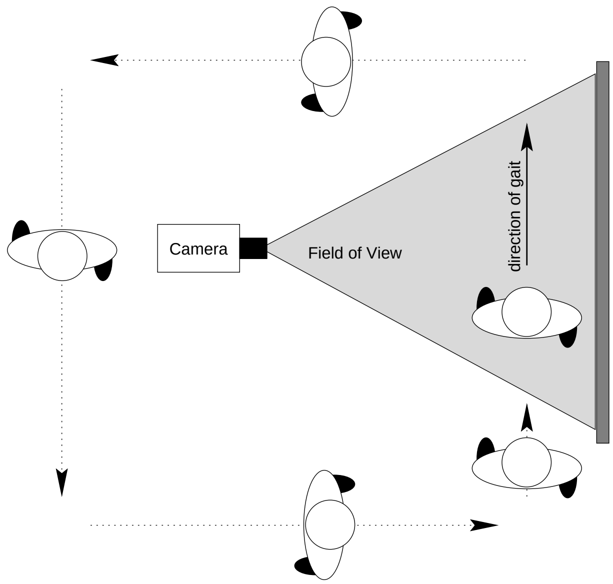


Figure 2: Schematic of experimental apparatus, plan view, for testing recognition using shape-of-motion features.

MLD can be considered to be a sampled version of the dense flow. If the sampled flow can produce the same or similar phase features, then shape-of-motion interpretation of gait sequences gives a model for non-structural gait perception.

4 Experiment

In order to determine if non-structural interpretation of MLDs is possible using shape-of-motion, we start with the gray-scale sequences acquired for the recognition experiments mentioned in the previous section. Figure 2 shows the experimental apparatus used to acquire that data. Subjects walked in a circular path around a fixed camera aimed at a static background. Only one subject was in the field of view at any time. The subjects walked for about 20 minutes and were recorded on video tape. Later a set of 25 sequences were digitized from the video tape representing the five subjects. Figure 3 shows an example of the images acquired, image number 39 of 78 in a sequence for one of the pedestrians. The sequences were processed by the system depicted in Figure 1.

We produced a set of MLDs, equivalent to the gray-scale sequences, by manually recording the coordinates of a set of joint positions for each frame in each sequence, and then

Figure 3: Sample image, number 39 from a sequence of 78, of an experimental subject walking across the field-of-view of the camera. Images were initially full-color, 640 by 480 *pixels*. They were resampled and cropped to 320 by 160 *pixels*.

creating synthetic MLDs from the lists of coordinates. This allows us to see the relationship between shape-of-motion features for the two forms of sequences. A custom annotation tool, written for X Windows, allows us to record the coordinates. The tool displays images in a sequence one at a time. While any single image is displayed, mouse clicks on the image cause the image coordinates of the mouse position to be recorded in a file. The result of annotating a sequence is a set of files, one file for each image in the sequence, where each file contains a list of joint coordinates.

The joints we selected for annotation were toe, ankle and knee on left and right legs, left hip, wrist and elbow on right and left arms, left shoulder, neck and head. It is obvious that, because of occlusions, not all of the joints are visible in any frame. Therefore, although we recorded the joints in the same order for each image, missing joints because of occlusions mean that there is no correspondence recorded in the data.

The manual annotation process was both difficult and tedious. The difficulty arose primarily from the fact that clothing concealed the true joint positions. Among the more difficult joints to deal with were knees concealed within baggy pants, elbows concealed by dark long sleeve shirts, and hips. The result is that these joints could not be reliably identified and there is considerable variation in the data due to the person performing the annotation guessing at the joint position. Added to this problem is the fact that the errors in the joint positions vary cyclically with the gait. This occurs because a joint such as a knee is easy to identify under clothing when the limb is bent but difficult when the limb is straight. Therefore, annotation errors are dependent on the point in the gait cycle. The effect of the errors can be likened to the physical situation where the lights are loosely fixed to the subject and jiggle about as the subject walks. Only sequences for three of the original five subjects were annotated.

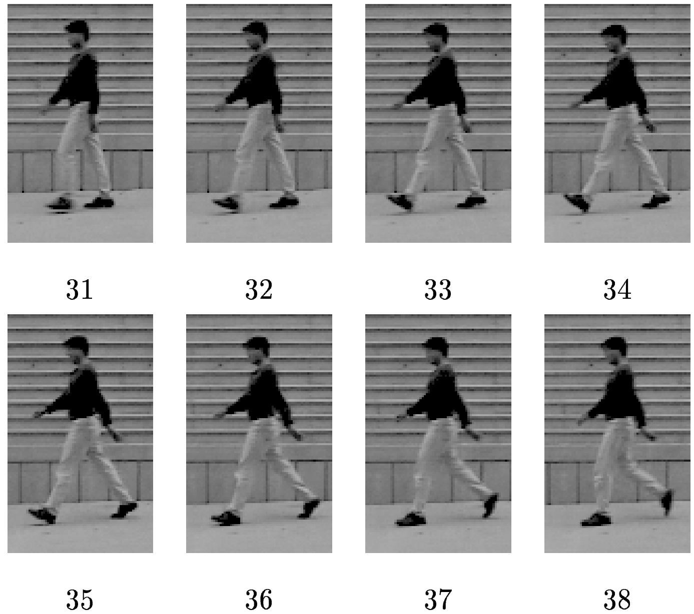
From the lists of joint coordinates, we synthesized MLDs by creating images that are black and then adding bright spots at the joint coordinates. The spots consist of a circle, radius 1.25 *pixels* of maximum brightness, 255, superimposed on a larger circle, radius 2.25 *pixels*, of half brightness, 128. Figure 4 shows eight sample frames extracted from a sequence for one of the three subjects, and the corresponding synthetic MLD.

Once the MLD sequence is created, we apply the procedure described in Figure 1. That is, we treat the MLD like any other image sequence, giving no special consideration to the fact that it is an MLD. Use of an optical flow algorithm not based on image gradients makes this possible. The dense optical flow becomes a field of zero flow with spots of flow corresponding to the lights that are moving. Occasionally the optical flow algorithm produces flow that does not represent true motion as joints pass in and out of occlusion. This occurs because the optical flow is not tracking points and is oblivious to the meaning of the dots. Errors caused by occlusions are treated like any other flow measured by the algorithm and receive no special consideration.

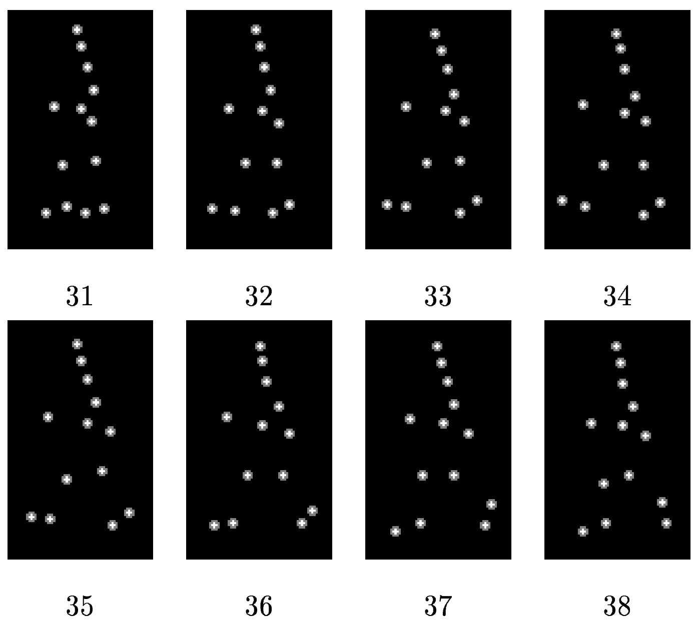
5 Results

We computed the optical flow using the algorithm of Bulthoff, Little and Poggio [5]. The algorithm computes the sum-of-absolute-differences over patches of two images, allowing the flow to be computed with different amounts of spatial support. We can express the spatial support in terms of σ for the Gaussian filter that weights the sum over a patch. Large spatial support yields coarse resolution flow while smaller support gives finer resolution. Coarse flow has fewer spurious errors than fine flow, but has less detail. To be sure that amount of spatial support of the flow computation was not a significant factor we computed the optical flow for the 15 MLD sequences at two different resolutions, one fine, $\sigma = 0.5$, and one coarse, $\sigma = 1.5$.

Since we are interested to see if the shape-of-motion features are preserved in the MLD, we compare the features generated from the MLDs with those generated for the gray-scale



(a)



(b)

Figure 4: Sample frames from a gait sequence: (a) gray-scale sequence and (b) corresponding frames from the manually derived MLD. Images are for subject number 2, sequence 3, images 31 to 38 of 78.

sequences in earlier experiments. In those earlier experiments, the focus was on recognition and the goal was to find significant variations between subjects in the experiment. Here we compare the mean values of features over the entire sample. This does not lead to conclusions about recognition, but indicates if the features change when going from the grey scale images to MLDs.

Figure 5 shows plots comparing the phase features for MLDs and grey-scale sequences for two resolutions. The error bars shown represent two standard deviations from the mean value. x_c, x_{wc}, x_d, y_d, a_c, a_{wc}, x_{uwc}, and y_{vwc} match well with respect to the variations in these values, and in consideration of the errors inherent in the manual annotation of light positions. Close inspection of the scalar signals indicated that a_d was mostly noise for MLDs, and therefore should be discarded. We came to this conclusion because the a_c signal appeared to be nearly identical to a_{wc}. Therefore, a_d, the difference of the two signals, should be zero, leaving just noise. The large error bars and differences between coarse and fine flow for a_d tend to support this. Only y_{wc} and y_{uwc} have poor matches for both resolutions. x_{vwc} matches poorly for the fine resolution flow, but matches slightly better at the coarser resolution.

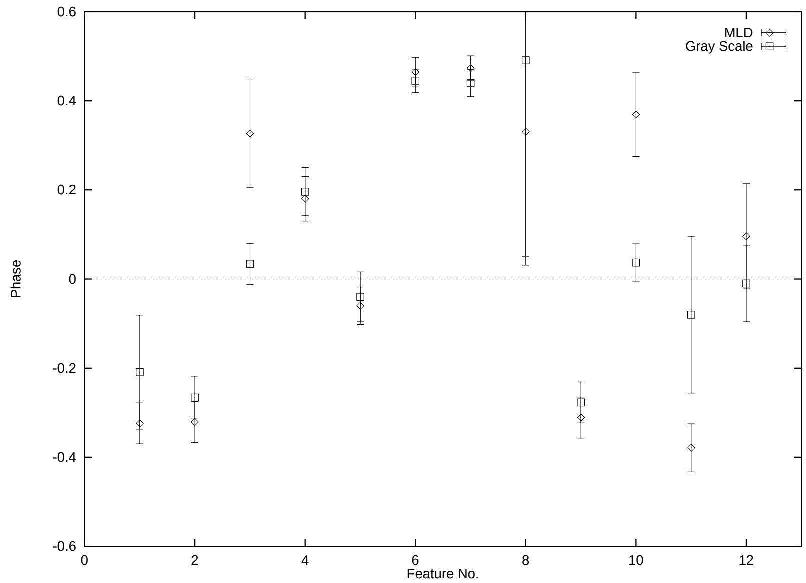
The results indicate that, as far as shape-of-motion features are concerned, the distribution of the optical flow is captured by the flow in the MLD. This supports the assertion that MLDs may be thought of as point sampling of a flow distribution.

6 Discussion

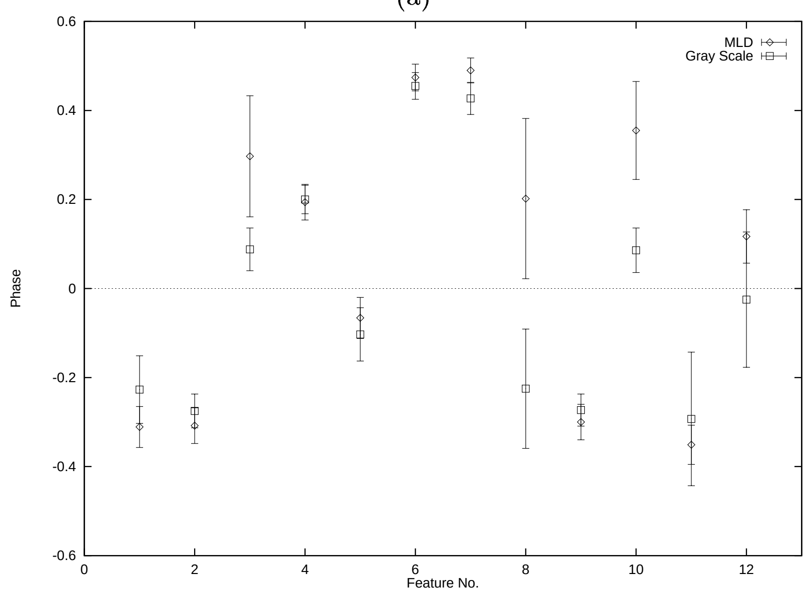
6.1 Psychophysical Observations

Perhaps the most important and fundamental psychophysical result for MLD gait perception is that of Johansson [7]. Johansson observed that the viewers of MLDs rarely failed to recognize correctly that the lights represented a person walking. Johansson assumed a structural explanation for this observation.

Results presented here show that shape-of-motion features are mostly preserved in MLDs. Since our previous work has shown the value of these features in characterizing a gait and allowing recognition, these new results suggest a non-structural explanation for gait recognition from MLDs. Although the shape-of-motion features do vary between indi-



(a)



(b)

Figure 5: Plots comparing the shape-of-motion features computed from gray scale images and from MLDs: (a) fine spatial flow and (b) coarse flow. Feature numbers correspond to features in the following order: (1) x_c, (2) x_{wc}, (3) y_{wc}, (4) x_d, (5) y_d, (6) a_c, (7) a_{wc}, (8) a_d, (9) x_{uwc}, (10) y_{uwc}, (11) x_{vwc}, and (12) y_{vwc}. Phase values are normalized to lie in the interval $[-0.5, 0.5]$.

viduals, they are consistent enough over all individuals to allow discrimination between gaits and other types of motions. Shape–of–motion features are not structural and are preserved in MLDs, so non-structural perception of gait from MLDs is possible using shape-of-motion features.

Barclay *et al.* [1] showed that a detectable amount of gender recognition is possible from MLDs, although the recognition was only slightly better than chance. We did not test for recognition of gender specifically. Of the three subjects, one was female. This is not nearly a large enough sample from which to draw reasonable conclusions. Ideally for recognition, the error bars in Figure 5 would be large while error bars for individuals would be small. Previous results for gray-scale sequences showed excellent recognition of individuals but our results with the smaller MLD data set were inconclusive. Recognition appeared to be successful, better than chance, using certain features. However, we hesitate to draw firm conclusions because the most significant feature for recognition turned out to be a_d which we believe to be mostly noise for MLDs. It is possible that the recognition was random. A larger data set is necessary to be able to draw reasonable conclusions about recognition. We mention recognition here because shape-of-motion features may still provide an explanation for the psychophysical results, especially in light of the fact that the MLD recognition of gender by humans is weak too.

Sumi looked at the effect of inverting the display of an MLD on recognition. His observations showed that while most people were often able to recognize a gait from an inverted MLD, they reported that the gait looked odd and failed to recognize that it was, in fact, normal but inverted. While we have not tested this explicitly, a complete knowledge of the model enables us to predict what will happen to phase features if the MLD is inverted. Inverting the MLD will reverse the phase of some, but not all of the features. For example, any phases of y-coordinates of centroids are reversed, i.e., rotated by $180°$. Simultaneously playing the sequences backwards in time, as Sumi did, alters the phase of all signals by reversing the signs of all relative phases. That is, if signal A leads B by $25°$, then after time reversal B leads A by $25°$. Thus under y-axis and time inversion, we expect phases to be altered in a systematic way. Some phase relationships would be preserved due to being twice inverted, while others would not. This may explain the common, but not universal, perception reported by Sumi, i.e., the perception of a human gait. Also, it seems likely that a human observer, after recognizing a gait using non-structural means, will try to map the expected human structure onto the pattern of dots. Sumi's subjects reported seeing gaits but thought that the gait was odd. Perhaps the oddness represents a failure to map human structure onto the stimulus when it is believed that the stimulus represents a human.

Bertenthal and Pinto present ideas about motion perception that are perhaps closest to our own. We certainly agree that phase is an essential part of gait recognition, but we differ in our ideas about the features for which the phase is important. While Bertenthal and Pinto look at the phase of the pendular motion of limbs, we look at the phase of the shape-of-motion features. However, the two are certainly related and this is highlighted by the following analysis.

Consider the image of a pedestrian to be a collection of moving points. Although the points move together in a pattern with the gait, we can treat the points separately. As a simple approximation, the motion for each point in the pedestrian can be expressed as the sum of a linear motion and an oscillatory motion. For example, let the x-coordinate of an arbitrary point i be

$$x_i(t) = x_{i0} + v_x t + A_i \cos(\omega t + \phi_i), \tag{1}$$

where x_{i0} is a constant, v_x is the mean velocity of the person, and A_i, ω and ϕ_i are the amplitude, frequency and phase of the oscillation. $x_{i0} + v_x t$ is the linear part of the motion. All points share the same frequency, ω, but vary in A_i and ϕ_i depending on where they are in the body. The x-coordinate of the centroid is

$$\bar{x}(t) = \frac{1}{n}\sum_i x_i(t) = \frac{1}{n}\sum_i \{x_{i0} + v_x t + A_i \cos(\omega t + \phi_i)\}, \tag{2}$$

where n is the number of image points in the pedestrian. As part of extracting a phase feature we discard the linear portion of the motion leaving

$$\bar{x}(t) = \frac{1}{n}\sum_i A_i \cos(\omega t + \phi_i). \tag{3}$$

The summation in Equation (3) is the sum of a set of phase vectors, or phasors. Phase vectors are commonly used to perform computations with rotating vectors that share a common frequency, such as in electrical power systems. In short, the summation can be treated as the sum of a set of vectors, each vector having magnitude A_i and direction ϕ_i.

The above leads to several conclusions. First, shape-of-motion features, at least those that are centroids, should be nearly the same for MLDs as they are for gray-scale data. The MLD merely restricts the set of points on the pedestrian that are available to view. For a proper image, the vector sum is over all the moving points. For an MLD the sum is over the moving points that are visible, i.e., those at the lights. The degree to which the shape-of-motion features are preserved in the MLD depends on how well the points chosen as lights in the display sample the full set of moving points. Second, if this model is correct, it is not necessary for the moving lights to be placed at joints. It is only necessary that they sample the motion adequately. Joints are an obvious place because they have the greatest amplitude of oscillatory motion on a limb, i.e., the knee moves more than the thigh and the ankle moves more than the shin, and so on. However, since there is no structure, there is no structural requirement to use joints. Finally, Bertenthal and Pinto's observations of the effect of joint phases can be explained with this model. A change in phase of the pendular motion of a limb becomes a change in the phase of a subset of the moving points for the pedestrian. Therefore, the vector sum is altered because a subset of the vectors being added is altered. The degree to which the sum is effected of course depends on the size of the limb and amount of phase perturbation. Note that the choice of points to illuminate for the MLD can affect the phase too. For example, many psychophysical experiments mark only the ankle and not the toe while we marked both. This in effect weights the summation of Equation (3) to place more emphasis on the movements of the feet.

6.2 Future Experiments

The experiment described in this paper merely shows that the same phase features available for recognition from gray scale sequences are mostly available in MLDs. If we wish to test the assertions about psychophysical results, we need to perform further experiments.

It should be possible to generate some sort of abstract pattern of dots or other shapes that varies in time to produce a sequence with the shape-of-motion characteristics of a pedestrian. Having produced such a sequence, we could have subjects of an experiment observe the pattern and report their perception. If successful, the observers would recognize a gait in the absence of any structural information. This would not only indicate a non-structural perception, but also support the validity of shape-of-motion phase features.

Results of such and experiment would also be useful in motion synthesis because they would indicate what features are necessary for realistic motion.

It may be interesting in such an experiment to have observers identify specific structural points in the gait they perceive (if they do in fact perceive a gait). The results may provide some insight as to where structure begins to play a part in perception. It may be that subjects will report the same perception observed by Sumi, i.e., they see gait, but something appears odd.

7 Conclusions

Shape-of-motion features have previously been shown to be useful for recognizing gaits from gray-scale sequences, without using kinematic structure. The experiment presented here indicates that these features are mostly preserved in MLDs. The implication is that interpretation of motion using shape-of-motion can be applied to MLDs, suggesting non-structural explanations of psychophysical observations.

This leads to the modeling of motion as a set of moving points that we characterize by a summation. Changes to points in the set change the summation. For the oscillatory signals of a gait, the change in sum can result in a change in phase and thus an altered perception. MLDs can, therefore, be viewed as a sample of the set of moving points. Perception is preserved in an MLD insofar as the MLD is a representative sample of the set.

References

[1] C. D. Barclay, J. E. Cutting, and Lynn T. Kozlowski. Temporal and spatial factors in gait perception that influence gender recognition. *Perception and Psychophysics*, 23(2):145–152, 1978.

[2] A. M. Baumberg and D. C. Hogg. Learning flexible models from image sequences. Technical Report 93.36, University of Leeds School of Computer Studies, October 1993.

[3] A. M. Baumberg and D. C. Hogg. Learning spatiotemporal models from training examples. Technical Report 95.9, University of Leeds School of Computer Studies, March 1995.

[4] B. I. Bertenthal and J. Pinto. Complementary processes in the perception and production of human movements. In L. B. Smith and E. Thelen, editors, *A Dynamic Systems Approach to Development: Applications*, pages 209–239. MIT Press, Cambridge, MA, 1993.

[5] H. Bulthoff, J. J. Little, and T. Poggio. A parallel algorithm for real-time computation of optical flow. *Nature*, 337:549–553, February 1989.

[6] C. Cedras and M. Shah. Motion-based recognition: a survey. *Image and Vision Computing*, 13(2):129–155, March 1995.

[7] G. Johansson. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14(2):201–211, 1973.

[8] G. Johansson. Visual motion perception. *Scientific American*, pages 76–88, June 1975.

[9] S. X. Ju, M. J. Black, and Y. Yacoob. Cardboard people: a parameterized model of atriculated image motion. In *2nd International Conference on Automatic Face and Gesture Recognition*, pages 38–44, Killington, Vermont, October 1996.

[10] J. J. Little and J. E. Boyd. Describing motion for recognition. In *IEEE Symposium on Computer Vision*, pages 235–240, November 1995.

[11] J. J. Little and J. E. Boyd. Recognizing people by their gait: the shape of motion. submitted to Videre, December 1996.

[12] R. Polana and R. Nelson. Detecting activities. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2–7, 1993.

[13] R. Polana and R. Nelson. Recognition of nonrigid motion. In *1994 DARPA Image Understanding Workshop*, pages 1219–1224, 1994.

[14] R. Polana and R. Nelson. Nonparametric recognition of nonrigid motion. Technical report, University of Rochester, 1995.

[15] S. Sumi. Upside-down presentation of the johansson moving light-spot pattern. *Perception*, 13:283–286, 1984.