Linguistic Descriptions for an Object in Motion

O. Sjahputera, P. Matsakis, J. M. Keller, R. Bondugula
Department of Computer Engineering & Computer Science
University of Missouri-Columbia, Columbia, MO 65211
{osb1e@mizzou.edu}

Abstract – **In ongoing work on spatial relations and scene interpretation, we present a system that linguistically describes the motion of an object in a temporal sequence. This description, called the dynamic linguistic description, is inferred from a sequence of static linguistic descriptions explaining the relative position, at different instances, between a moving object and a stationary object. In this preliminary work, the moving object is assumed to be moving in a straight path at a constant velocity. The scene is monitored from a fixed pose with a constant frame rate. The proposed system is potentially useful as a low-bandwidth remote observation system capable of linguistically reporting relative position and motion in a scene.**

Keywords: spatial relation, linguistic description of motion, scene interpretation

I. INTRODUCTION

Spatial relations play an important role in scene description. Various methods to assess spatial relations between objects in images have been proposed. For instance, Krishnapuram et al. [1] and Miyajima and Ralescu [2] used the *angle histogram* method. Matsakis and Wendling [3] introduced the concept of *force histograms* built on a solid theoretical foundation. They also showed that the histogram of forces generalizes and supersedes the histogram of angles.

Generating linguistic descriptions for spatial relations is another important task. Keller and Wang [4,5] used angle histograms in addition to some metric features as inputs to a system of neural networks that generated directional relationships between two objects. The resulting relationships were fed to a fuzzy rule-base system to produce a linguistic scene description. However, the linguistic terms used were coarse and this method failed to satisfy some criteria such as the semantic inverse principle [6]. In more recent work, Matsakis et al. [7] used two types of force histograms, constant forces (F_0) and gravitational forces (F_2). Features that represent directional relations between objects were extracted from each histogram. Heuristics rules were employed to combine the features which were fed into a fuzzy rule-base. This produced a linguistic description using much richer language that could be tailored to meet a user's needs.

In robotics, Skubic et al. [8] used force histograms to generate multi-level linguistic spatial descriptions of the surrounding environment based on sensor readings on the robot. The use of a few linguistic terms to describe the surroundings reduces the bandwidth required to maintain steady communication between the remotely operating robot and its human supervisor.

The task of detecting and capturing the path of a moving object in a scene has been widely discussed. Sun and Tan [9] presented a scene monitoring system with a static camera that estimated the background (stationary) as a two-dimensional site environment map. The foreground containing moving objects is treated as 3-D data carrying both spatial and temporal information. A scene monitoring system is not restricted to a static observation platform. Medioni et al. [10] proposed a system that analyzes the behavior of moving objects in a scene observed from a camera mounted on a moving platform. Using optical flow, regions that contain moving objects are extracted and attribute graph representations are used to infer their trajectories. Motion tracking has also been used in robotics. Luo and Chen [11] introduced a grey-fuzzy controller (GFC) that enables robots to track and follow the object in motion without requiring *a priori* knowledge on the dynamic models of the target and the tracker.

The system proposed in this paper produces linguistic descriptions that capture the direction of a moving object by tracking its relative positions with respect to a stationary object. The relative position at each time instant is represented by a *static* linguistic description (*s*) discussed in [6]. This creates a time-dependent sequence of observation. The sequence is processed based on the temporal ordering and the duration of contiguous observations for each *s*. We estimate the direction of motion and generate the *dynamic* linguistic description (*m*) to explain it. Other work on motion detection and representation rely on optical-flow-based features. Our system relies solely on the sequence of *s*. In this preliminary work, the moving object is assumed to move in a straight path at a constant velocity and observation is made from a fixed pose at a constant frame rate. This paper is organized as follows: Section II briefly discusses *s*, Section III describes the method to estimate the direction of motion and the method for generating *m*, Section IV contains experimental results, conclusions are given in section V.

II. STATIC LINGUISTIC DESCRIPTION

A static linguistic description (*s*) captures the relative position of an argument object (A) with respect to a reference object (B) at a given time. An *s* is produced using methods proposed in [7] that are based on the concept of the histogram of forces in [3]. Both histograms of gravitational forces (F_2) and that of constant forces (F_0) are calculated from the object pair (A,B). Both F_0 and F_2 represent the relative position of A with respect to B. A number of features are extracted from these histograms. These features represent the opinion of each histogram about the proposition *"A is in the direction δ of B"* where δ is one of the four primitive directions RIGHT, ABOVE, LEFT, BELOW. The opinions of F_0 and F_2 are combined to give us a compromised degree of truth for the proposition *"A is in the direction δ of B"*. The combined opinions are fed into a fuzzy rulebase, which determines the appropriate linguistic hedges for the *primary direction* (δ_1) and *secondary direction* (δ_2). The linguistic hedges are defined in a dictionary, which can be tailored to a user's needs. An example of *s* is *"A is {mostly} to the right of B, but {somewhat} above"*. Here, δ_1=*RIGHT*, δ_2=*ABOVE*, *"mostly"* and *"somewhat"* are the linguistic hedges for the primary and secondary direction respectively. *s* can also indicate a compound direction description, such as *"A is above-right of B"*. In [7], Matsakis et al. divided the plane into 24 conical regions, numbered from ⓪ to ㉓. When both objects A and B can be assimilated to points, each of these region's, *r*, corresponds to a specific primary and secondary linguistic hedge as shown in Fig. 1. *s* is then classified as a *regular* static description and denoted by L_r. If objects A and B cannot be assimilated to points (i.e. they are too close together), the *s* generated may contain additional hedges. We denote this type of descriptions as the *non-regular* descriptions. Fig. 2 shows two examples of this type of descriptions. Here, *s* for group 1 (three large buildings in black as the reference) and group 2 is *"Group 2 is {loosely} above-left of group 1"*. The *s* for group 1 and object 3 is *"Group 3 is {perfectly} above group 1, but {slightly} shifted to the right"*.

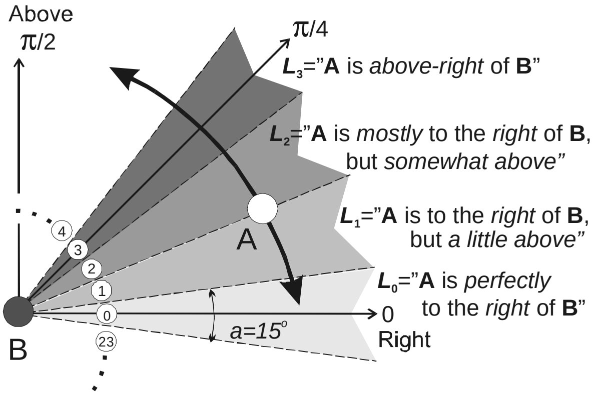

Fig. 1 The framework associating regular static linguistic descriptions L_r to a specific region of relative positions when both objects A and B are point-like objects.

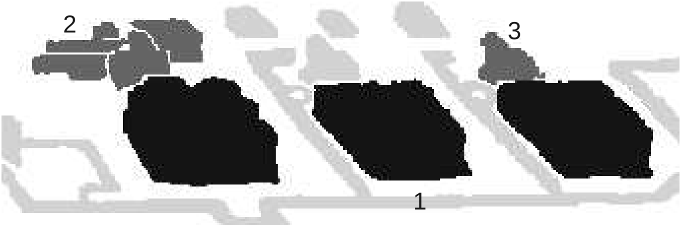

Fig. 2 Configurations involving objects that cannot be assimilated to points.

III. Dynamic Linguistic Description

III.1 Background

We are to observe a sequence of image frames containing objects A and B. Both objects can usually be assimilated to points, relatively small in size, and neither touching nor overlapping. Object A is assumed to move following a straight path with constant velocity, while object B is stationary. Hence, our input should contain regular *s*, L_r, only. However, non-regular *s* may also appear as an indication of a transition between two adjacent regions, and cases where objects A and B get too close. For instance, when object A crosses from region ① to ②, the description *"A is to the right of B, but {somewhat} above"* may appear during the transition from a regular description L_1, *"A is to the right of B, but {a little} above"*, to its adjacent regular description L_2, *"A is {mostly} to the right of B, but {somewhat} above"*.

When object A crosses from region ⓪ to ①, its L_r is expected to change from *"A is {perfectly} to the right of B"* to *"A is to the right of B, but {a little} above"*. Based on our knowledge of the framework shown in Fig. 1, we can infer that object A must be moving in a direction γ where 7.5°< γ <187.5°. A wide range of γ makes it difficult for us to come up with a reasonable linguistic description for the direction of motion. Information on the order of past and current L_r alone is insufficient for this purpose. Another piece of information available is the *duration of observation* for each state of static linguistic description. Assuming constant object speed and observer frame rate, the number of observations can be used to represent the relative *distance* traveled by object A to go through a region, as illustrated in Fig. 3.

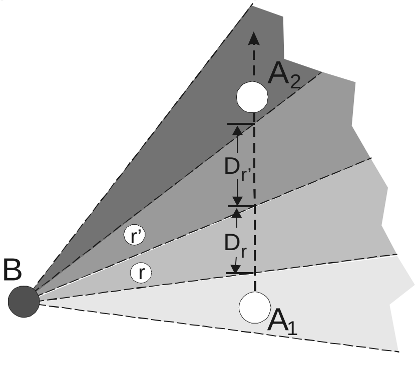

Fig. 3 Object A moves from A1 to A2. The numbers of observations D_r and $D_{r'}$ can be used as estimates for the distances object A traveled through region *r* and *r'* respectively.

III.2 Estimating Direction of Motion

The full sequence of static linguistic descriptions is denoted as: $S=\{s_0,s_1,\cdots,s_t,\cdots,s_T\}$. Ideally, *S* contains only *regular* descriptions, L_r, like those shown in Fig. 1. However, as mentioned in section III.1, non-regular descriptions may also appear, especially when objects A and B are getting too close. Let L_r be the *regular* static linguistic description associated with region *r*. The duration of observation for L_r is defined as

$D_r=|\{t \mid s_t=L_r,\ s_t\in S\}|$. *Complete-observation* of L_r is achieved if L_r is preceded by $L_{r''}$ and succeeded by $L_{r'}$ in S where both $L_{r'}$ and $L_{r''}$ are *regular* static linguistic descriptions associated to the regions adjacent to the r. In other words, we must have observed L_r fully from the beginning, starting when s_t changes to L_r till the end when s_t changes to $L_{r'}$. Illustration of D_r as a distance estimate is shown in Fig. 3. Once two complete observations have been achieved consecutively, say L_r and $L_{r'}$, we estimate the direction γ of motion:

$$\gamma=\tan^{-1}\left[\left(\frac{N-1}{N+1}\right)\cot(a)\right]+\left(r+\frac{1}{2}\right)a \tag{1}$$

a=15° is the opening angle for the conical regions, $N=D_r/D_{r'}$ if region r is closer to the primary direction δ_1, $N = D_{r'}/D_r$ otherwise. Fig. 4 shows the framework for γ.

III.3 Linguistic Terms of Motion description

To produce the dynamic linguistic description m, we use a similar framework as in [7] where the general direction is divided into ranges. Let L_r and $L_{r'}$ be the two regular static descriptions used to calculate γ. Here, γ outlines the straight path of the motion, but it does not indicate which way the motion is going along the path. Fig. 4 shows that the same γ can be used to represent object A moving either "*right-upward*" or "*left-downward*". We can view γ as a periodic value in $[-\pi/2, \pi/2]$ whose range is represented by a half-circle centered around a line perpendicular to δ_1 as shown in Fig. 4. The perpendicular line represents the direction γ=0°. If γ is outside this range, say γ=95°, then we represent γ using its opposite equal, in this case γ=-85° which is within the range.

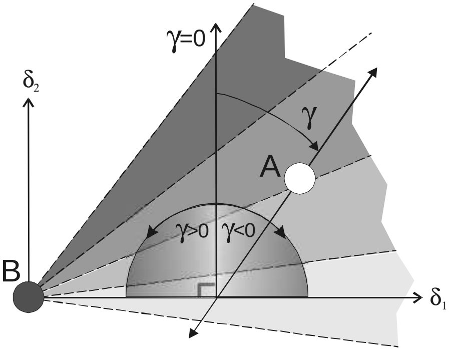

Fig. 4 The framework for γ outlining the estimated path for object A.

To detect which way the object A is moving, we need to consider the temporal ordering of the observed static linguistic descriptions. Let r and r' be the two adjacent regions whose regular static descriptions are used to calculate γ. If region r is closer to δ_1, we denote this as a ⊖ transition where relative position of object A from B is getting "closer" to δ_1. On the other hand, if r is closer to δ_1 then we denote this as a ⊕ transition (the relative position of object A is shifting away from δ_1). Let γ_1 and γ_2 be the primary and secondary directions of the motion. The values for γ_1 and γ_2 are determined from δ_1 and δ_2 using the rules given in Table 1.

Table 1 Rules for determining the primary and secondary direction of motion, γ_1 and γ_2, from the primary and secondary direction of the static linguistic descriptions, δ_1 and δ_2.

γ	**⊕ transition**		**⊖ transition**	
	γ_1	**γ_2**	**γ_1**	**γ_2**
$\pi/4>\gamma\geq 0$	δ_2	$\delta_1+\pi$	$\delta_2+\pi$	δ_1
$\pi/2\geq\gamma\geq\pi/4$	$\delta_1+\pi$	δ_2	δ_1	$\delta_2+\pi$
$0>\gamma>-\pi/4$	δ_2	δ_1	$\delta_2+\pi$	$\delta_1+\pi$
$-\pi/4\geq\gamma>-\pi/2$	δ_1	δ_2	$\delta_1+\pi$	$\delta_2+\pi$

The linguistic terms RIGHT, UPWARD, LEFT, and DOWNWARD are used for γ_1 and γ_2. We have three linguistic hedges for γ_1 and two for γ_2. For compound directions, we use the term "*diagonally*" followed by both γ_1 and γ_2. A set of crisp rules is used to assign these hedges based on the value of $\phi=\min(|\gamma|, \pi/2-|\gamma|)$. ϕ is the angle of motion viewed from γ_1 and γ_2 perspective. When object A is moving exactly to direction γ_1 we obtain ϕ=0. The rule set is shown in Table 2.

Table 2 Rules for generating linguistic hedges for dynamic linguistic description. The angle a for the conical regions is 15**°.**

		γ_1	**γ_2**
ϕ	$a/2\geq\phi\geq 0$	***No hedge***	***No secondary description***
	$3a/2\geq\phi>a/2$	"Primarily"	"A little"
	$5a/2\geq\phi>3a/2$	"Mostly"	"Somewhat"
	$\pi/4\geq\phi>5a/2$	**Compound Direction**	

III.4 Non-Regular Static Descriptions

During a transition from L_r to $L_{r'}$, we may observe a sequence of non-regular s_t with duration $\tilde{D}_r$. We distribute the evidence from the non-regular s_t by modifying $N=(D_r+0.5\tilde{D}_r)/(D_{r'}+0.5\tilde{D}_r)$ for ⊕ transitions, and $N=(D_{r'}+0.5\tilde{D}_r)/(D_r+0.5\tilde{D}_r)$ for ⊖ transitions.

III.5 Dynamic Descriptions for Non-Complete Observations

If our observation starts with L_0 (region ⓪), D_1 and D_2 are achieved when $L_1\rightarrow L_2$ and $L_2\rightarrow L_3$ respectively, allowing the use of Eq. 1 to obtain γ. Prior to obtaining D_1 and D_2, we are still able to generate some dynamic descriptions for the motion at transition points $L_0\rightarrow L_1$ and $L_1\rightarrow L_2$. Consider the configuration given in Fig. 5 where object A is moving from position A_1 to A_2. The transition $L_0\rightarrow L_1$ indicates that object A has moved from region ⓪ to region ①. The range of all possible directions that makes such crossing possible is defined as $]\alpha_1,\alpha_2[$, α_m is the median. We can simply generate dynamic descriptions based on α_1, α_2, and α_m to give the user a general direction where object A is moving to. Similarly, the range $]\beta_1, \beta_2[$ with β_m as the median is established when object A moves from region ① to region ② ($L_1\rightarrow L_2$). We can narrow the range by taking the intersection of the two ranges as shown in Fig. 5 resulting in a narrower range delimited by $]\beta_1,\alpha_2[$ with $(\beta_m+\alpha_m)/2$ as the new median.

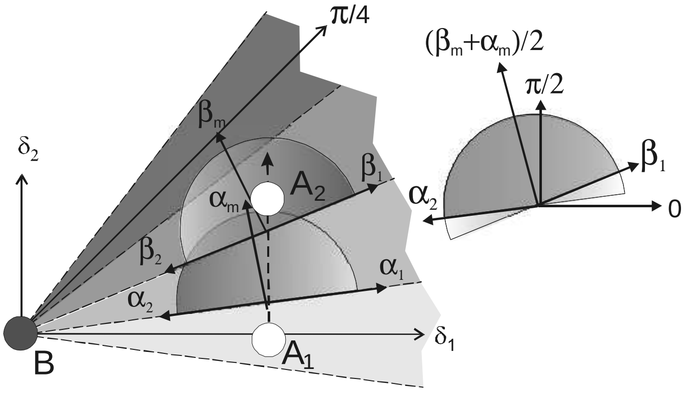

Fig. 5 Framework for establishing ranges of possible motion directions involving non-complete observation.

The angles {α_1, α_2, α_m}, and {β_1, β_2, β_m} are measured from δ_1 toward δ_2, with δ_1 as the zero reference. To generate dynamic linguistic description for each angle we need to determine γ_1 and γ_2 for each α_k and β_k, k={1,2,m}. The map given in Fig. 6 shows how to determine γ_1 and γ_2 from δ_1 and δ_2. For example, if δ_1=RIGHT, δ_2=ABOVE, and α_m=97.5° then γ_1=UPWARD and γ_2=LEFT. To determine linguistic hedges for γ_1 and γ_2 in this example, define $\phi=\min(|\alpha_m-\gamma_1|,|\alpha_m-\gamma_2|)$ which gives us ϕ=7.5°. Using the rules in Table 2 and ϕ we just obtained, the dynamic description for α_m would be *"Object A is moving upward"*. Similarly, we come up with the descriptions *"Object A is moving to the left"* and *"Object A is moving {primarily} to the right, but {a little} upward"* for α_1 and α_2, respectively. These descriptions allow us to linguistically describe the range $]\alpha_1,\alpha_2[$ with α_m as the median.

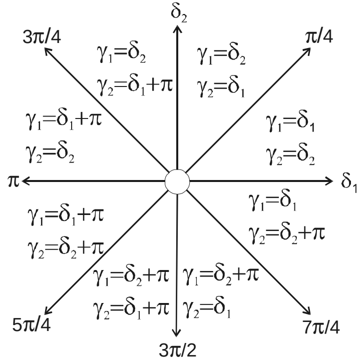

Fig. 6 Transforming δ_1 and δ_2 in static description to γ_1 and γ_2 for dynamic description in non-complete observation cases.

IV. EXPERIMENTAL RESULTS

We have implemented a simulation program for the proposed method using the C language with OpenGL graphic library and Glut interface. The program has two output windows: a graphic window to display both objects and a text window to display static and dynamic linguistic descriptions. Both objects are circular in shape and represented using 12-vertex polygons. The radius for each object is determined individually. The user also determines the coordinates for object B and the starting and ending coordinates that define the path for object A. The path is traced using the Bresenham's line algorithm.

In Fig. 7, object B is placed at (150,200), and object A travels from (250,10) to (250,390), a perfectly vertical path upward. A white vertical line marks the path. Object radius is 10 pixels wide. The expected dynamic linguistic description is *"A is moving upward"*. Smaller circles along the path mark the spots where transitions of static linguistic descriptions occur and a dynamic linguistic description is generated. The circles are called *transition points* (TP), and are numbered in the order of occurrence (1 to 8).

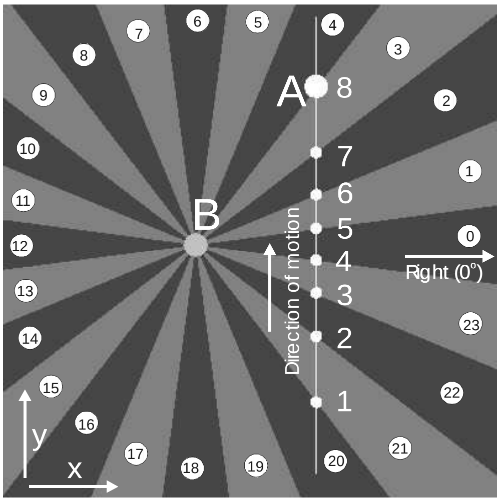

Fig. 7 Experiment for *"Object A is moving upward"*.

The linguistic descriptions generated at TP-1 to 8 are shown in Table 3. At TP-1 and TP-2, the dynamic descriptions generated are based on non-complete observations, because the number of *complete observations* L_r available is still less than two. At TP-1, object A crosses from ⑳ to ㉑. However, D_{20} does not necessarily represent the complete path across ⑳. By the time object A crosses from ㉑ to ㉒ at TP-2, the only *complete observation* L_r we have achieved is L_{21}. The dynamic description generated at TP-2 is a little more specific than that of TP-1. However, both are still too vague to give us a reasonably good estimate where object A is going. Eq. 1 can be used starting at TP-3 onwards, where two consecutive *complete observations* L_r are available. In the case of TP-3, the two *complete observations* L_r are L_{21} and L_{22} with δ_1=RIGHT δ_2=BELOW, and $N=D_{22}/D_{21}$. From this experiment, D_{22}=35.5 and D_{21}=55, and Eq. 1 returns γ=-1.3°. Since this is a ⊖ transition, using transformation rules given in Table 1, we have γ_1=UPWARD and γ_2=LEFT. The rules in Table 2 return no hedge for γ_1 and no secondary direction description is available. Hence, the dynamic description is *"Object A is moving upward"*, which correctly describes the direction of object A. The same dynamic description is generated at each TP onwards.

Table 3 Static and dynamic linguistic descriptions from Fig. 7.

	Static & Dynamic Descriptions
L_{20}	A is mostly below B, but somewhat to the right.
1	Object A is likely to move in the general direction of: to the right-upward. Possible range of direction extends from the above direction up to: downward-to the right. Or, extends up to: mostly upward, but somewhat to the left
L_{21}	A is below-right of B.
2	Object A is likely to move in the general direction of: to the right-upward. Possible range of direction extends from the above direction up to: upward-to the left. Or, extends up to: mostly to the right, but somewhat downward
L_{22}	A is mostly to the right of B, but somewhat below.
3	Object A is moving upward.
L_{23}	A is to the right of B, but a little below.
4	Object A is moving upward.
L_0	A is perfectly to the right of B.
5	Object A is moving upward.
L_1	A is to the right of B but a little above.
6	Object A is moving upward.
L_2	A is mostly to the right of B, but somewhat above.
7	Object A is moving upward.
L_3	A is above-right of B.
8	Object A is moving upward.
L_4	A is mostly above B but somewhat to the right

For the second experiment, we place object B at (150, 250), and object A moves from (10,10) to (390,390) following a diagonal path as shown in Fig. 8. At TP-3, we start applying Eq. 1 since L_{17} and L_{18}, are *complete-observation*. Here, δ_1=BELOW and δ_2=LEFT, D_{17}=48.5 and D_{18}=27.5, and Eq. 1 returns γ=-45.8°. This is a ⊝ transition, hence γ_1=UPWARD, γ_2=RIGHT, and ϕ=44.2° giving us a *compound direction*. The dynamic linguistic description for TP-3 is *"Object A is moving diagonally to the right and upward"*. Similarly, a dynamic description is generated at each transition points onwards. Here, all dynamic descriptions (Table 4) describe the direction of motion of object A correctly.

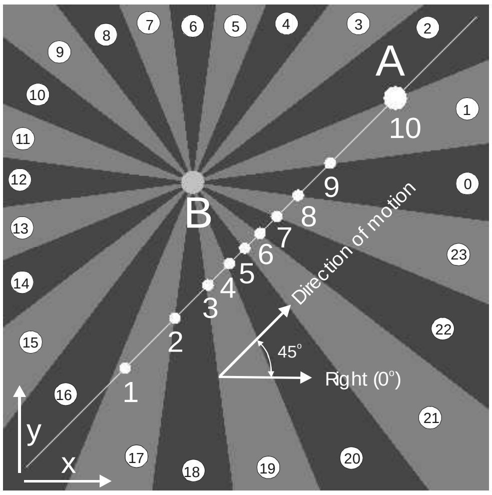

Fig. 8 Experiment for *"Object A is moving diagonally to the right and upward"*.

Table 4 Static and dynamic linguistic descriptions from Fig. 8.

	Static & Dynamic Descriptions
L_{16}	A is mostly below B, but somewhat to the left.
1	Object A is likely to move in the general direction of: mostly to the right, but somewhat downward. Possible range of direction extends from the above direction up to: mostly upward, but somewhat to the right. Or, extends up to: primarily downward, but slightly to the left
L_{17}	A is below B, but a little to the left
2	Object A is likely to move in the general direction of: primarily to the right, but slightly downward. Possible range of direction extends from the above direction up to: mostly upward, but somewhat to the right. Or, extends up to: downward.
L_{18}	A is perfectly below B
3	Object A is moving diagonally to the right and upward.
L_{19}	A is below B, but a little to the right
4	Object A is moving diagonally to the right and upward.
L_{20}	A is mostly below B, but somewhat to the right
5	Object A is moving diagonally to the right and upward.
L_{21}	A is below-right of B
6	Object A is moving diagonally to the right and upward.
L_{22}	A is mostly to the right of B, but somewhat below
7	Object A is moving diagonally to the right and upward.
L_{23}	A is to the right of B, but a little below
8	Object A is moving diagonally to the right and upward.
L_0	A is perfectly to the right of B
9	Object A is moving diagonally to the right and upward.
L_1	A is to the right of B, but a little above
10	Object A is moving diagonally to the right and upward.
L_2	A is mostly to the right of B, but somewhat above

In the third example, we move object B to (175,225), closer to the path of object A as shown in Fig. 9. The results given Table 5 show that for the most part our system is able to generate accurate description of the motion. However, as object A gets too close to object B (TP-5 to 10), the dynamic descriptions show an increasing degree of inaccuracies. The reason is that the two objects cannot really be assimilated to points (ie., one assumption that this work is based on does not hold). At TP- 5, the dynamic description is *"Object A is moving mostly to the right, but somewhat upward"* rather than the expected *"Object A is moving diagonally to the right and upward"*. However, this is not too far off since the hedges used, *"mostly"* and *"somewhat"*, represent the range of ϕ adjacent to the correct range for this motion. Upon closer inspection, we find an increase of non-regular static descriptions during the transition from region ⑱ to ⑲. In this case the description is *"A is perfectly below B, but slightly shifted to the right relative to B"*. This increase causes a delay during the transition from region ⑱ to ⑲, where TP-4 is detected well inside region ⑲ rather than at the border between ⑱ and ⑲. The effects of non-regular static descriptions are observed at each transition point from TP-5 to TP-10, representing the section of the path where object A is at the closest distance from object B. However, the system still manages to generate a respectably accurate dynamic description during this time, until the distance between the two objects opens up and the correct dynamic descriptions begin to reappear at TP-11 onwards.

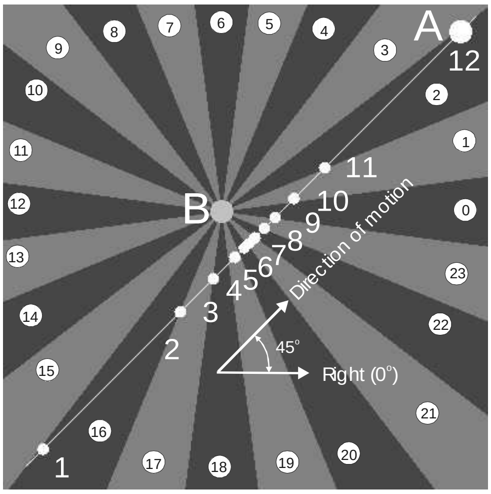

Fig. 9 Experiment for *"Object A is moving diagonally to the right and upward"* with object B positioned closer to the path.

Table 5 Static and dynamic linguistic descriptions from Fig. 8.

	Static & Dynamic Descriptions
L_{15}	A is below-left of B.
1	Object A is likely to move in the general direction of: to the right-downward. Possible range of direction extends from the above direction up to: upward-to the right. Or, extends up to: mostly downward, but somewhat to the left
L_{16}	A is mostly below B, but somewhat to the left.
2	Object A is likely to move in the general direction of: mostly to the right, but somewhat downward. Possible range of direction extends from the above direction up to: upward-to the right. Or, extends up to: primarily downward, but slightly to the left.
L_{17}	A is below B, but a little to the left
3	Object A is moving diagonally to the right and upward.
L_{18}	A is perfectly below B.
4	Object A is moving diagonally to the right and upward
L_{19}	A is below B, but a little to the right
5	Object A is moving mostly to the right, but somewhat upward
L_{20}	A is mostly below B, but somewhat to the right.
6	Object A is moving mostly to the right, but somewhat upward.
L_{21}	A is below-right of B
7	Object A is moving mostly to the right, but somewhat upward.
L_{22}	A is mostly to the right of B, but somewhat below.
8	Object A is moving mostly upward, but somewhat to the right.
L_{23}	A is to the right of B, but a little below.
9	Object A is moving mostly upward, but somewhat to the right
L_0	A is perfectly to the right of B
10	Object A is moving mostly upward, but somewhat to the right
L_1	A is to the right of B, but a little above.
11	Object A is moving diagonally upward and to the right.
L_2	A is mostly to the right of B, but somewhat above.
12	Object A is moving diagonally upward and to the right.
L_3	A is above-right of B.

IV. CONCLUSIONS & FUTURE WORK

Using a simple object form and consistent motion behavior, we have demonstrated a method to linguistically describe the direction of motion of an object in a scene by simply observing the pattern of linguistic descriptions that explain the relative position of the moving object relative to a stationary object in each frame. One contribution offered by this work is the introduction of a new method that allows us to detect and describe motion in a scene without using the actual image data or other pixel-based features extracted from the image. The next step is to include linguistic distance descriptions as part of the input. The ability to detect and describe changes in direction of motion and velocity will be added soon. At this moment, each dynamic linguistic description is generated without taking into account past dynamic descriptions. A temporal fusion of the direction information will help us obtain a more robust system. Further improvements will allow our system to deal with non-point-like objects, non-straight line paths and uncertainties related with non-regular static linguistic descriptions.

ACKNOWLEDGEMENT

The Office of Naval Research (ONR) grant N00014-96-0439 supports this work.

REFERENCES

[1] R. Krishnapuram, J.M. Keller, Y. Ma, "Quantitative Analysis of Properties and Spatial Relations of Fuzzy Image Regions," *IEEE Trans. Fuzzy Systems*, vol. 1, no. 3, p.222–233, 1993.

[2] K. Miyajima, A. Ralescu, "Spatial organization in 2-D segmented images: Representation and recognition of primitive spatial relations," *Fuzzy Sets & Syst.*, vol. 65, no. 2/3, pp. 225-236, 1994.

[3] P. Matsakis, L. Wendling, "A new way to represent the relative position between areal objects," *IEEE Trans. PAMI*, vol. 21, pp. 634-643, July 1999.

[4] X. Wang, J.M. Keller, "Human-based spatial relationship generalization through neural/fuzzy approaches," *Fuzzy Sets & Syst.*, vol 101, no. 1, pp.5-20, 1999.

[5] J.M. Keller, X. Wang, "A fuzzy rule-based approach for scene description involving spatial relationships," *Comput. Vis. Image Understand.*, vol. 80, pp.21-41, 2000.

[6] J. Freeman, "The modeling of spatial relations," *Comput. Graph.* Image Process., vol. 4, pp. 156-171, 1975.

[7] P. Matsakis, J.M. Keller, J. Marjamaa, O. Sjahputera, "Linguistic description of relative positions in images," *IEEE Trans. SMC-B*, vol. 31, no. 4, pp. 573-588, August 2001.

[8] M. Skubic, G. Chronis, P. Matsakis, J. Keller, "Generating Linguistic Spatial Descriptions from Sonar Readings Using the Histogram of Forces," *IEEE Int. Conf. on Robotics and Automation, Proc.*, May 2001, pp. 485-90.

[9] H. Sun, T. Feng, T. Tan, "Spatio-temporal segmentation for video surveillance," *Pattern Recognition, Proc. 15th Intl. Conf.*, vol. 1, pp. 843–846, 2000.

[10] G. Medioni, I. Cohen, F. Bremond, S. Hongeng, R. Nevatia, **"**Event detection and analysis from video streams,*"* *IEEE Trans. PAMI*, vol. 23, no. 8, pp. 873-889, Aug. 2001.

[11] R.C. Luo, T.M. Chen, K.L. Su, "Target tracking using a hierarchical grey-fuzzy motion decision-making method," *IEEE Trans. SMC-A*, vol. 31, no. 3, pp. 179–186, May 2001