Figure/Ground Segregation from Human Cues

Artur M. Arsenio
MIT Computer Science and Artificial Intelligence Laboratory
Cambridge, MA 02139, USA
Email: arsenio@csail.mit.edu

***Abstract*— This paper presents a new embodied approach for object segmentation by a humanoid robot. It relies on interactions with a human teacher that drives the robot through the process of segmenting objects from arbitrarily complex, non-static images. Objects from a large spectrum of different scenarios were successfully segmented by the proposed algorithms.**

I. Introduction

Embodied vision [3] extends far behind the opportunities crated by active vision systems [1], [4]. The human (and/or robot) body is used not only to facilitate perception, but also to change the world context so that it is easily understood by the robotic creature (Cog, the humanoid robot used throughout this work, is shown in Figure 1 through different segmentation scenarios).

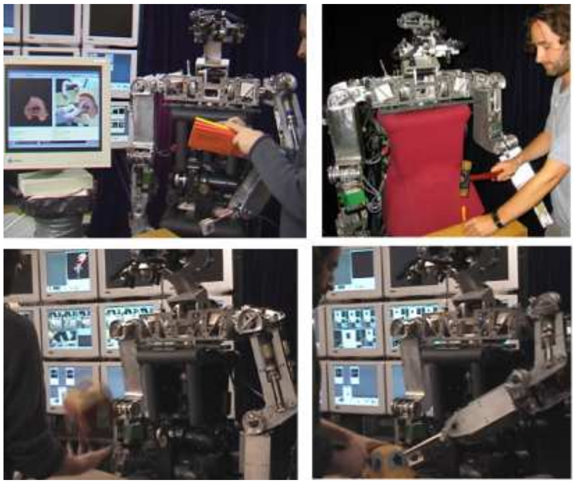

Fig. 1. Cog, the humanoid robot used throughout this work, shown through several learning scenarios. These images correspond to real experiments from which objects were separated from the background.

Embodied vision methods will be demonstrated with the goal of simplifying visual processing. This is achieved by selectively attending to the human actuator (*Hand*, *Arm* or *Finger*), or the robot actuator. Indeed, primates have specific brain areas to process the hand visual appearance [14].

Focus will be placed on a fundamental problem in computer vision - *Object Segmentation* - which will be dealt with by detecting and interpreting natural human/robot task behavior such as waving, shaking, poking, grabbing/dropping or throwing objects. Object segmentation is truly a key ability worth investing effort so that other capabilities, such as object/function recognition can be developed.

A. Embodied object segmentation

The number of visual segmentation techniques is vast. An active segmentation technique developed recently [7] relies on poking objects with a robot actuator. This strategy operates on first-person perspectives of the world: the robot watching its own motion. However, it is not suitable for segmenting objects based on external cues. Among previous object segmentation techniques it should be stressed the minimum-cut algorithm [13]. Although a good tool, it suffers from several problems which affect non-embodied techniques. Indeed, object segmentation on unstructured, non-static, noisy, low resolution (128×128 images) and real-time images is a hard problem:

- ▷ object with similar color/texture as background
- ▷ multiple moving objects in a scene
- ▷ robustness to luminosity variations

Segmentations must also present robustness to variations in world structure. Mobility constraints (important for heavy objects) poses additional difficulties, since motion cannot be used to facilitate the segregation process.

Distinguishing an object from its surroundings – the figure/ground segregation problem – will be dealt by exploiting shared world perspectives between a cooperative human and a robot. We argue for a visual embodied strategy for object segmentation, which is not limited to active robotic heads. Instead, embodiment of an agent is exploited by probing the world with a human/robot arm. This strategy proves not only useful to segment movable objects, but also to segment object descriptions from books, as well as large, stationary objects (such as a table) from monocular images.

This paper is organized as follows. The next three sections describe different protocols a human instructor might use to boost the robot's object segmentation capabilities (the overall algorithmic control structure is shown in Figure 2). Segmentation by demonstration is described in Section II. This technique is especially well suited for segmentation of fixed or heavy objects in a scene, such as a table or a drawer, or objects drawn or printed in books. Section III presents object segmentation through active object actuation. Objects are waved or shaken by a human actor in front of the robot. Objects that are difficult to wave but easily acted on (for instance, by poking them) are segmented as described in Section IV. Experimental object segmentation results are presented in each section. Finally, Section V draws the conclusions.

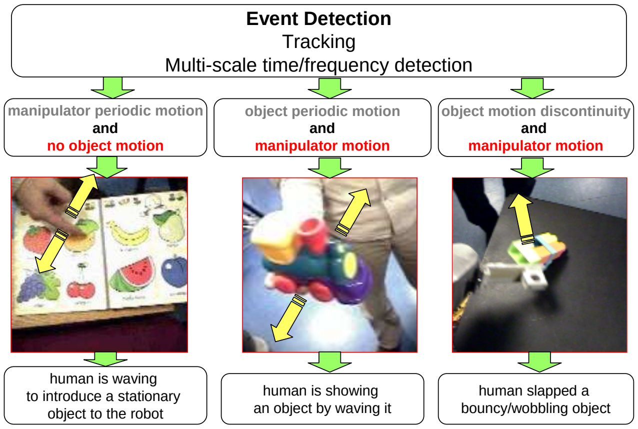

Fig. 2. Image objects are segmented differently according to scene context. The selection of the appropriate method is done automatically. After detecting an event and determining the trajectory of periodic points, the algorithm determines whether objects or actuators are present, and switches to the appropriate segmentation method.

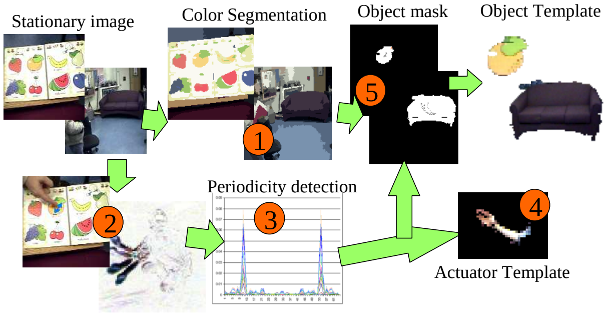

Fig. 3. A standard color segmentation algorithm computes a compact cover for the image. The actuator's periodic trajectory is used to extract the object's compact cover – a collection of color cluster sets.

II. Segmentation by Demonstration

We propose a human aided object segmentation algorithm to tackle the figure-ground problem. Indeed, a significant amount of contextual information may be extracted from a periodically moving actuator. This can be framed as the problem of estimating $p(o_n|v_{B_{\vec{p},\epsilon}}, act^{per}_{\vec{p},S})$, the probability of finding object n given a set of local, stationary features v on a neighborhood ball B of radius ϵ centered on location p, and a periodic actuator on such neighborhood with trajectory points in the set $S \subseteq B$. The following algorithm implements the estimation process to solve this figure-ground separation problem (see Figure 3):

1) A standard color segmentation [6] algorithm is applied to a stationary image (stationary over a sequence of consecutive frames)
2) A human actor waves an arm on top of the object to be segmented
3) The motion of skin-tone pixels is tracked over a time interval (using the Lucas-Kanade Pyramidal algorithm), and the energy per frequency content is determined for each point's trajectory
4) Periodic, skin-tone points are grouped together into the arm mask [2].
5) The trajectory of the arm's endpoint describes an algebraic variety [9] over N^2. The target object's template is then given by the union of all bounded subsets (the color regions of the stationary image) which intersect this variety

An affine flow-model is estimated (using a least squares minimization criterium) from the optical flow data, and used to determine the trajectory of the arm/hand/finger position over the temporal sequence. Periodic detection is then applied at multiple scales. Indeed, for an arm oscillating during a short period of time, the movement might not appear periodic at a coarser scale, but appear as such at a finer scale. If a strong periodicity is not found at a larger scale, the window size is halved and the procedure is repeated again for each half.

The algorithm consists of grouping together the colors that form an object. This grouping works by having periodic trajectory points being used as seed pixels. The algorithm fills the regions of the color segmented image whose pixel values are closer to the seed pixel values, using a 8-connectivity strategy. Therefore, points taken from waving are used to both select and group a set of segmented regions into the full object. Clusters grouped by a single trajectory might either form or not form the smallest compact cover which includes the full object (depending on intersecting or not all the clusters that form the object). After the detection of two or more temporally and spatially closed trajectories this problem vanishes.

A. *Perceptual Organization*

According to Gestalt psychologists, the whole is different than the sum of its parts – the whole is more structured than just a group of separate particles. The technique we suggest segregates objects from the background without processing local features such as *textons* or contours [12]. The proposed grouping paradigm differs from Gestalt grouping rules for perceptual organization. These rules specify how parts are grouped for forming wholes, and some of them are indeed exploited by our grouping method: *Similarity* and *Proximity* rules are embedded on the color segmentation algorithm; moving periodic points in an image sequence are also grouped together. It is worthy to stress that this technique solves very easily the *figure and ground* illusion (usually experienced when gazing at the illustration of a white vase on a black background – the white vase is segregated just by having an human actor waving on (or the black faces, if the human selects them).

B. *Experimental Results*

Embodiment of an agent is exploited for probing the world with a human arm. The segmentation by demonstration algorithm is then applied to segment the visual appearance of objects from the background. This algorithm was used to segment both object descriptions from books and large, stationary objects (such as a table) from monocular images. Figure 4 shows segmentations for a random sample of object segmentations (furniture items), together with statistical results for such objects.

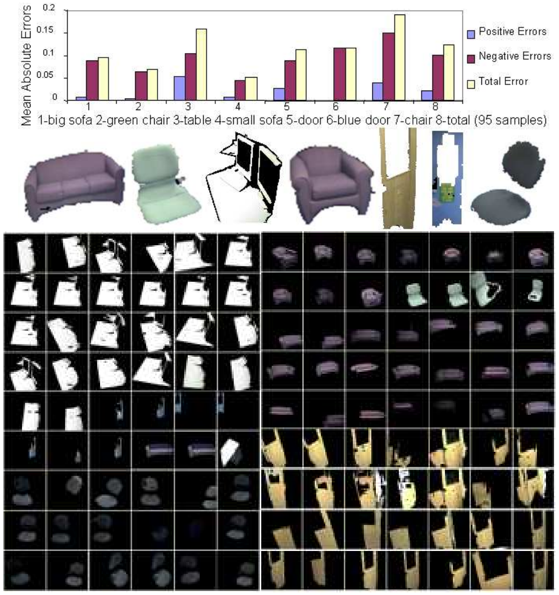

Fig. 4. Statistics for the furniture items (a set of segmentation samples is also shown). Errors are given by (template area - object's real visual appearance area)/(real area). Positive errors stand solely for templates with larger area than the real area, while negative errors stand for the inverse. Total errors stand for both errors. The real area values were determined manually. A chair is grouped from two disconnect regions by merging temporally and spatially close segmentations.

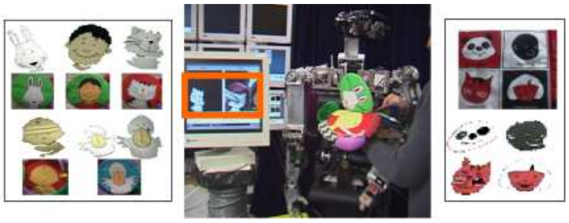

Fig. 5. Teaching the visual appearance of objects to a robot. Left illustration shows segmentation results – on top – of book pages – on bottom – from a book made of fabric (middle illustration). Right illustration shows multiple segmentations acquired from a page of another fabric book. It is worthy to stress that segmentation on books made of fabric textile poses additional difficulties, since pages deform easily, creating perspective deformations, shadows and object occlusions.

This strategy relies heavily in human-robot interactions. It is essential to have a human in the loop to introduce objects from a book to the robot (as a human caregiver does to a child), by tapping on their book's representations – figure 5 shows a human teaching a robot from a fabric book, by tapping on relevant figures. Indeed, human caregivers often use books to introduce a child to a diverse set of (in)animate objects, exposing the latter to an outside world of colors, forms, shapes and contrasts, that otherwise might not be available to a child (such as the image of a whale). Since this learning aid helps to expand the child's knowledge of the world, it is a useful tool for introducing new informative percepts to a robot (whales do not appear often on research labs!). To corroborate this argument, the human-centered segmentation algorithm was successfully applied to extract templates for animals (including a whale), clothes, fruits, geometric shapes, musical instruments and other elements from books (as shown in Figure 6), under varying light conditions (no environment setup).

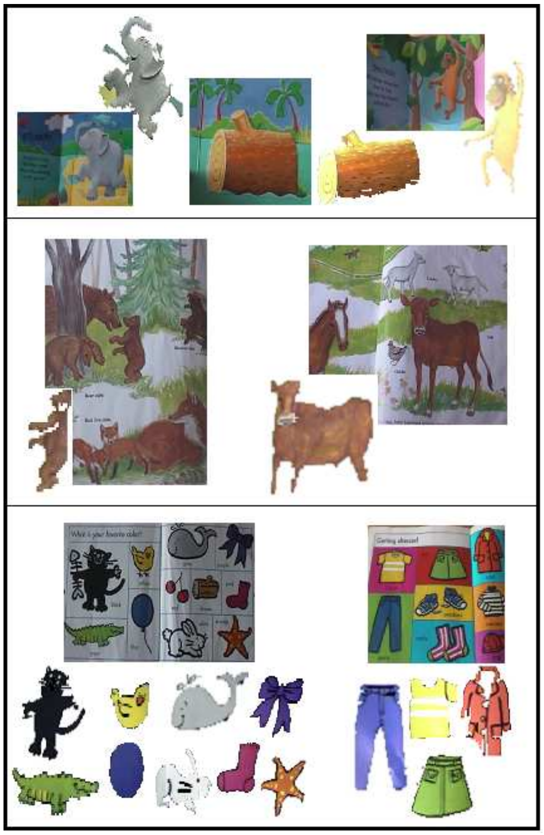

Fig. 6. Figure/ground segregation from paper books (books targeted for infants and toddlers). Templates for several categories of objects (for which a representative sample is shown), were extracted from dozens of books. The book pages shown were not recorded simultaneously as the segmentations, but they are useful to identify the background from which the object's template was extracted. (top) Clusters of colors were grouped together into an elephant, a piece of wood and a monkey. (middle) A bear and a cow are segmented from a book. These two animals are the union of a large set of local color clusters. (bottom) Segmentation of several elements from books.

Typical errors might arise from objects with similar color to their background, for which no perfect differentiation is possible, since the intersection of the object's compact cover of color regions with the object's complementary background is not empty (see Figure 7). High color variability within an object create additional grouping difficulties (the object's compact cover of color clusters contains too many sets – harder to group).

a) Visual Illusions:: It is worth stressing that this grouping technique solves very easily the *figure and ground* illusion. This is usually experienced when gazing at the illustration of a white vase in a black background – the white vase (or the black faces) is segregated just by having a human actor waving above it (see figure 8).

C. Perceptual Grouping of Spectral Cues

Another important property of objects is the texture of their surfaces – and texture has complementary properties to color. Texture is closely related to the distribution both

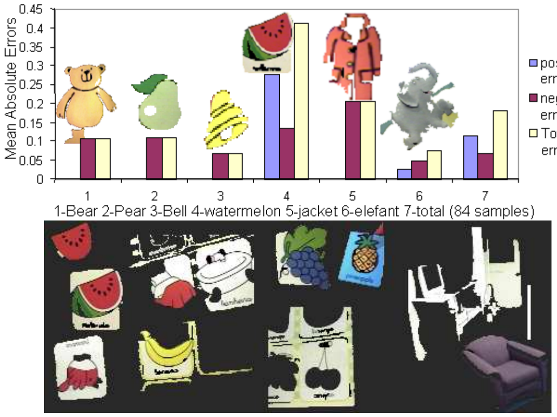

Fig. 7. (top) Statistical analysis for segmentation errors from books. (bottom) A representative set of templates illustrating sources of errors. The watermelon, banana, bed and white door have color clusters with identical color – white – to its background, for which no differentiation is possible, since the intersection of the object's compact cover of color regions with the background is not empty. High color variability of the sofa pixels create grouping difficulties (the compact cover contains too many sets – harder to group, unless object is well covered by human hand trajectories). The cherries reflect another source of errors - very small images of objects are hard to segment.

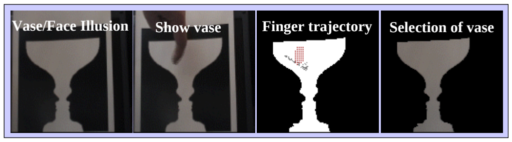

Fig. 8. Solving the vase-figure illusion. Having a human tapping on the vase extracts the vase percept, instead of the faces percept. The third image from the left shows both sampled points of a human finger and the finger's endpoint trajectory.

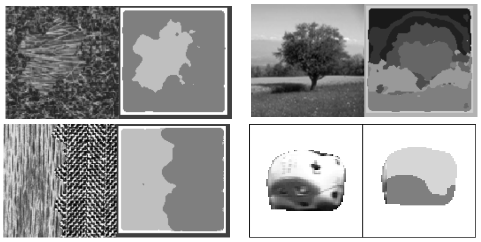

Fig. 9. Texture segmentation. The number of texture clusters was determined automatically.

in space and frequency of an object's appearance. Gabor filters and Wavelets are tools often applied to solve the texture segmentation problem.

Texture Segmentation: Objects templates will be segmented into regions of similar texture using a standard texture segmentation approach, as follows. A Wavelet transform is initially applied to the image to be segmented. This is approximately equivalent to a family of Gabor filters sampling the frequency domain in a log-polar manner. The original image is correlated with a Daubechies-4 filter using the Mallat pyramid algorithm, at two levels $(N = 2)$ of resolution [16], resulting in $n = 4 + 16 = 20$ coefficient images (see figure 9). All these coefficient images are then up-sampled to the original size, using a 5×5 gaussian window for interpolation. This way, more filtering is applied to the coefficients at the N^{th} level. For each pixel, the observation vector is then given by n wavelet coefficients. A mixture of gaussians is then applied to probabilistically model the input data by clusters. It is therefore necessary to learn the parameters for each cluster and the number of clusters. The former is estimated using the expectation-maximization (EM) algorithm [8]. The latter uses an agglomerative clustering approach based on the Rissanen order identification criteria [15]. The image is then segmented according to the cluster of gaussians: a point belongs to an image region if it occurs with maximum probability for the corresponding gaussian.

Grouping Textures: We can then apply a modified version of the perceptual grouping from human demonstration method to group perceptual (texture) cues, replacing the standard color segmentation algorithm applied to a stationary image by the texture segmentation algorithm. This approach is especially useful to segment objects with a homogeneous texture but heterogeneous colors (see figure 10).

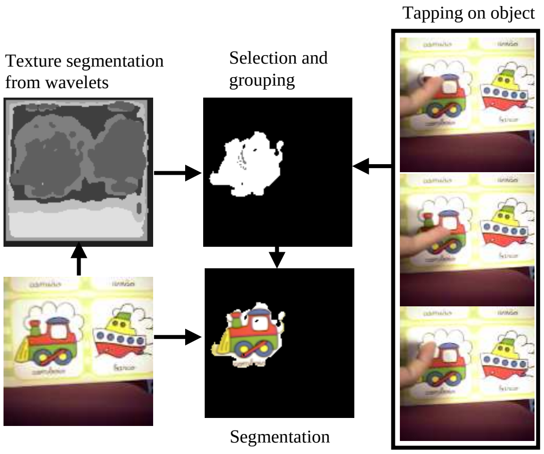

Fig. 10. Texture segmentation of an image and grouping of texture clusters. The method is applied successfully for a case in which the grouping for color regions fails, due to color similarity with background (the white color).

D. Improving Texture Segmentation Accuracy

A more advanced and complex texture segmentation algorithm – the normalized cuts algorithm [13] – can be applied for improved performance, to divide the images into texture clusters which might then be grouped together through human-robot interactions. Therefore, we demonstrate in figures 11 and 12 how to improve experimental segmentation results originally presented by [13], by grouping texture clusters into meaningful percepts of a church and the sky, and a mountain.

We have just seem in this section a strategy for a human to transmit to the robot information concerning the appearance of objects that cannot be moved relatively to their background. We now move forward to demonstrate perceptual grouping for objects that can be actively actuated.

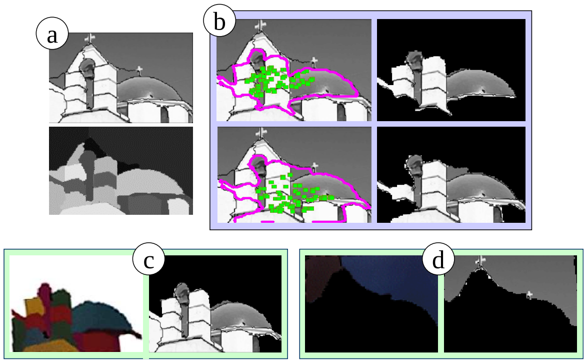

Fig. 11. Improving normalized cuts' segmentation results. Texture clusters segmented from an image (a) by a normalized cuts algorithm [13] are grouped together (b,c) into a church, and into the sky (d), by having a human showing the picture of the church to the robot and tapping on it.

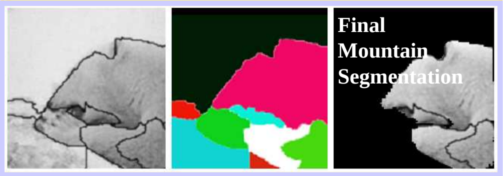

Fig. 12. Texture clusters segmented from a nature scene [13] are grouped together into a mountain through human demonstration.

III. Segmentation Driven by Active Actuation

This technique is triggered by the following condition: the majority of periodic points are generic in appearance, rather than drawn from the hand or finger. A visual scene might contain several moving objects, which may have similar colors or textures as the background. Multiple moving objects create ambiguous segmentations from motion, while large similarities between figure and background makes this figure/ground separation problem harder. However, a human teacher can facilitate robot's perception by waving or shaking an object in front of the robot, so that the motion of the object is used to segment it, as follows: Moving image points are initialized and tracked thereafter over a time interval; Their trajectory is then evaluated using a Short Time Fourier transform (STFT), and tested for a strong periodicity. Periodic, non-skin points are then grouped into a unified object.

A. Tracking

A grid of points homogeneously sampled from the image are initialized in the moving region, and thereafter tracked over a time interval of approximately 2 seconds (65 frames). At each frame, each point's velocity is computed together with the point's location in the next frame.

The motion trajectory for each point over this time interval was determined using four different methods. Two were based on the computation of the image optical flow field - the apparent motion of image brightness - and consisted of 1) the Horn and Schunk algorithm [10]; and 2) Proesmans's algorithm - essentially a multiscale, anisotropic diffusion variant of Horn and Schunk's algorithm. The other two algorithms rely on discrete point tracking: 1) block matching; and 2) the Lucas-Kanade pyramidal algorithm. The Lucas-Kanade algorithm achieved the best results.

B. Multi-scale Periodic Detection

A STFT is applied to each point's motion sequence,

$$I(t, f_t) = \sum_{t'=0}^{N-1} i(t')h(t'-t)e^{-j2\pi f_t t'} \qquad (1)$$

where h is usually a Hamming window, and N the number of frames. In this work a rectangular window was used. Although it spreads more the width of the peaks of energy than the Hamming window, it does not degrade overall performance, and decreases computational times.

Periodicity is estimated from a periodogram determined for all signals from the energy of the STFTs over the frequency spectrum. These periodograms are processed by a collection of narrow bandwidth band-pass filters. Periodicity is found if, compared to the maximum filter output, all remaining outputs are negligible. The periodic detection is applied at multiple time scales. If a strong periodicity is found, the points implicated are used as seeds for segmentation.

C. From Perceptual Grouping to Object Recognition

Now that periodic motion can be spatially detected and isolated, the waving behavior guides the segmentation process:

1) The set of moving, non-skin [5] points tracked over a time window is sparse. Hence, an affine flow-model is applied to the periodic flow data to recruit other points within uncertainty bounds
2) Clusters of points moving coherently are then covered by a non-convex polygon – approximated by the union of a collection of overlapping, locally convex polygons [2].

This algorithm is much faster than the minimum cut algorithm [13], and provides segmentations of similar quality to the active minimum cut approach presented by [7]. Figure 13 presents an error analysis together with a snapshot of the system running on the humanoid robot.

Our human-computer interactive approach introduces a humanoid robot to new percepts stored in its surrounding world. Such percepts must then be converted into an useful format through an object recognition scheme, which enables the robot to recognize an object in several contexts and under different perspective views. This object recognition algorithm needs to cluster objects by classes according to their identity. Such task was implemented through color histograms objects are classified based on the relative distribution of their color pixels. New object templates are classified according to their similarity with other object templates in an object database (Figure 13). A multi-target tracking algorithm (which tracks good features using the Lucas-Kanade Pyramidal algorithm) was developed to keep

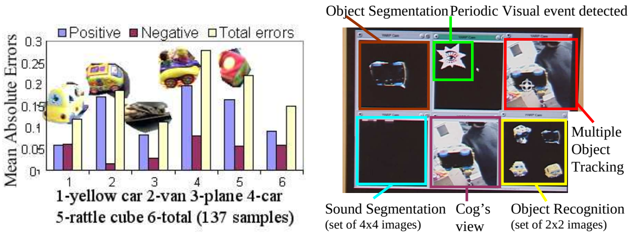

Fig. 13. (left) Statistical results for the segmentation of the objects shown. Results shown in graph (6) correspond to data averaged from a larger set of objects on the database (right) Visual segmentations are used to initialize a multi-target tracking algorithm, to keep track of the objects' positions. An object recognition algorithm, which matches templates based on color histograms, is also shown running.

track of object locations as the visual percepts change due to movement of the active head.

Figure 14 shows segmentation samples for a few objects under a variety of perspective deformations. This approach is robust to other scene objects and/or people moving in the background (they are ignored as long as their motion is non-periodic).

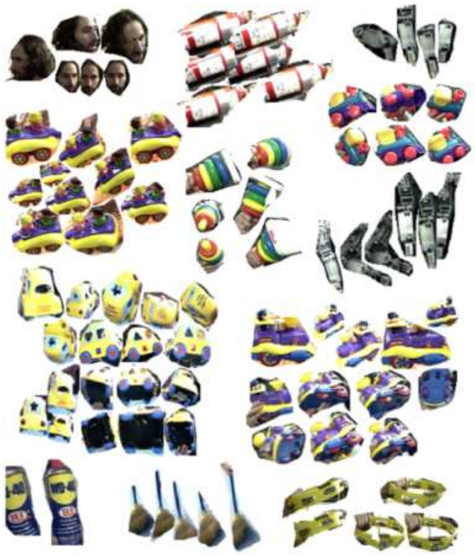

Fig. 14. Sample of objects segmented from oscillatory motions, under different views. Several segmentations of the robots arm (and upper arm) are also shown, obtained from rhythmic motions of robot's arm (and upper arm, respectively) in front of a mirror. Shown are sample segmentations from a large corpora consisting of tens of thousands of segmentations computed by the real-time segmentation algorithm.

D. Deformable Contours

An interesting trade-off results from the procedure of both tracking and segmenting. Tracking through discrete methods over large regions (necessary for fast motions) requires larger correlation windows.For highly textured regions, points near an object boundary are correctly tracked. But whenever an object moves in a textureless background, background points near the object's boundary will be tracked with the object, creating an artificial textureless membrane surrounding the object. Nonetheless, accurate segmentation is still possible using deformable contours, initialized to the boundaries of the object templates, and allowed to contract to the real boundaries. For textureless backgrounds, the deformable contour is attracted to the closest edges that define the object's boundary.

Snakes (active contour models) are deformable contours widely used, that move under the influence of image forces. We can define a deformable contour by constructing a suitable deformation energy $P_i(z)$, where z represents the contour points. The external forces on the contour result from a potential $P_e(z)$. A snake is a deformable contour that minimizes the energy [11]: $P(z) = P_e(z) + P_i(z)$, $z(s) = (x(s), y(s))$. For a simple snake, the internal deformation energy is:

$$\int_0^T \frac{\alpha(s)}{2}|\dot{z}(s)|^2 + \frac{\beta(s)}{2}|\ddot{z}(s)|^2 ds \tag{2}$$

where the differentiation is in respect to s. This energy function models the deformation of a stretchy, flexible contour $z(s)$ and includes two physical parameter functions: $\alpha(s)$ controls the tension and $\beta(s)$ the rigidity of the contour. To attract the snake to edge points we specify the external potential $P_e(z)$,

$$P_e(z) = \int_0^T P_{image}(z(s))ds \tag{3}$$

where $P_{image}(x, y)$ is a scalar potential function defined over the image plane [11]. The local minima of P_{image} are the snake attractors. Hence, the snake will have an affinity to intensity edges $|\nabla I(x, y)|$ computed by a Canny detector,

$$P_{image}(x, y) = -D(|\nabla I(x, y)|) \tag{4}$$

where D is the distance transform function, used to calculate the distance of image pixels to the deformable contour. The snake is interpolated at each step to keep the distance between each point constant. The snake achieves high convergence rates towards the minimum, improving therefore segmentation results, as shown by Figure 15.

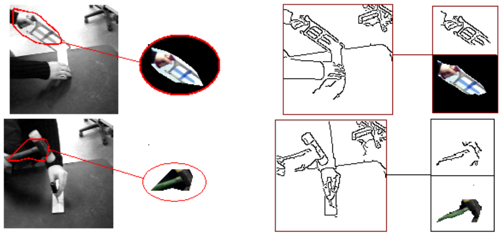

Fig. 15. Deformable contours improve quality for object segmentations. The contour, which is initialized to the original template's boundaries, is tracked towards the nearest edges, removing the undesirable *membrane*.

IV. Segmentation Through Discontinuities

The discontinuous motion induced on an object whenever a robot (or a human instructor) acts on it can be used for segmenting the object. In order to detect discontinuous events, an algorithm was developed to identify interactions among multiple objects in the image:

1) A motion mask is first derived by subtracting gaussian filtered versions of successive images and placing non-convex polygons around any motion found.
2) A region filling algorithm is applied to separate the mask into regions of disjoint polygons (using a 8-connectivity criterion).
3) Each of these regions is used to mask a contour image computed by a Canny edge detector.
4) The contour points are then tracked using the Lucas-Kanade pyramidal algorithm.
5) An affine model is built for each moving region from the position and velocity of the tracked points. Outliers are removed using the covariance estimate for such model.

Spatial events are then defined according to the type of objects' motion discontinuity [2]. This strategy was used to detect events such as grabbing, dropping, poking, assembling, disassembling or throwing objects, and to segment objects from such events by application of the grouping algorithm in Section III-C. A random sample of segmentations is presented in Figure 16. Figure 17 presents an analysis on segmentation quality, together with snapshots of the algorithm running on the humanoid robot.

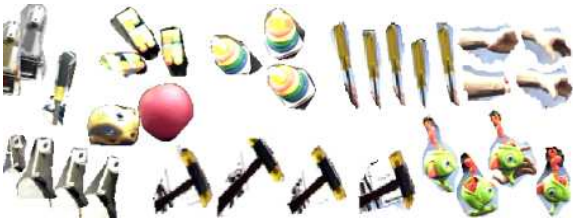

Fig. 16. Sample segmentations from object's discontinuous motions actuated by humans and the robot. Not only the object's visual appearance is segmented from images, but also the robot's end-effector appearance.

Errors								Total
Mean Error	0.24	0.03	0.25	1.21	0.14	0.29	0.17	0.48
Std	0.19	0.09	0.05	0.43	0.16	0.38	0.28	0.56
Mean abs error	0.29	0.07	0.25	1.21	0.20	0.37	0.30	0.52

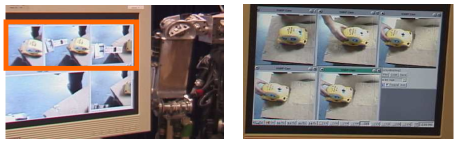

Fig. 17. (top) Statistical error analysis for the segmentation of the objects shown in the table (bottom) System running on the robot – left/right images show the detection of a poking/grabbing event created by the humanoid robot/human actor, respectively.

V. Conclusions

In this paper we introduced the human in the learning loop to facilitate robot perception. By exploiting movements with a strong periodic or discontinuity content, the robot's visual system segments a wide variety of objects from images, with varying conditions of luminosity and a different number of moving artifacts in the scene. The detection is carried out at different time scales for a better compromise between frequency and spatial resolution. The techniques presented can be used in a passive vision system (no robot is required), with a human instructor guiding the segmentation process. But a robot may also guide the segmentation process by himself, such as by poking.

We proposed a grouping strategy to segment objects that are not allowed to move and therefore might be difficult to separate from the background. This human-centered technique is especially powerful to segment fixed or heavy objects in a scene or to teach a robot through the use of books.

Through interactions of a robot with a human instructor, the latter facilitated the robot's segmentation task by providing additional grouping cues. Similarly, human teachers facilitate children's perception and learning during child development phases. It was also shown that the autonomous acquisition of such informative percepts is what it is need to train an object recognition algorithm.

Acknowledgements

Project funded by DARPA as part of the "Natural Tasking of Robots Based on Human Interaction Cues" under contract DABT 63-00-C-10102, and by the Nippon Telegraph and Telephone Corporation as part of the NTT/MIT Collaboration Agreement. Author supported by Portuguese grant PRAXIS XXI BD/15851/98.

References

[1] Aloimonos, J., Weiss, I. and Bandopadhay,A.: Active Vision. Int. Journal on Computer Vision (1987) 333–356
[2] Arsenio, A. M.: Embodied Vision - Perceiving Objects from Actions. IEEE Int. Workshop on Human-Robot Interactive Communication (2003)
[3] Arsenio, A. M.: Towards and Embodied and Situated AI. International FLAIRS conference (2004)
[4] Bajcsy, R.: Active perception. Proceedings of the IEEE, v.76, n.8, (1988) 996–1005
[5] Breazeal, C.: Sociable Machines: Expressive Social Exchange Between Humans and Robots. MIT PhD thesis, Cambridge, MA (2000)
[6] Comaniciu, D. and Meer, P.: Robust Analysis of Feature Spaces: Color Image Segmentation. IEEE Conference on Computer Vision and Pattern Recognition (1997)
[7] Paul Fitzpatrick, P.: From First Contact to Close Encounters: A Developmentally Deep Perceptual System for a Humanoid Robot, MIT PhD Thesis (2003)
[8] Gershenfeld, N.: The nature of mathematical modeling. Cambridge university press (1999)
[9] Harris, J.: Algebraic Geometry: A First Course (Graduate Texts in Mathematics, 133). Springer-Verlag, January (1994)
[10] Horn, B.: Robot Vision. MIT Press, (1986)
[11] Kass, M., Witkin, A. and Terzopoulos, D.: Snakes: Active Contour Models. International Journal of Computer Vision, (1987) 21–31
[12] Malik, J., Belongie, S., Shi, J. and Leung, T.: Textons, Contours and Regions: Cue Integration in Image Segmentation. IEEE Int. Conf. on Computer Vision (1999)
[13] Shi, J. and Malik, J.: Normalized cuts and image segmentation. IEEE Trans. on Pat. Anal. and Machine Intelligence, (2000) 888–905
[14] Perrett, D., Mistlin, A., Harries, M. and Chitty, A.: Understanding the visual appearance and consequence of hand action. Vision and action: the control of grasping. Ablex (1990) 163–180
[15] Rissanen, J.: A Universal Prior for Integers and Estimation by Minimum Description Length. Annals of Statistics (1983), 417–431
[16] Strang, G. and Nguyen, T.: Wavelets and Filter Banks. Wellesley-Cambridge Press (1996)