

Following the Flock

Distributed Formation Control with Omnidirectional Vision-Based Motion Segmentation and Visual Servoing

BY RENÉ VIDAL,
OMID SHAKERNIA,
AND SHANKAR SASTRY

Vision seems to be a critical component in animals' abilities to respond to their neighbors' motions. For example, flocks of birds and schools of fish are able to maintain a coherent formation without explicit communication among individuals. Our long-term goal involves enabling groups of mobile robots to visually maintain formations in the absence of communication (Figure 1). Towards that end, we consider a formation control scenario in which the followers use motion segmentation for estimating the image position of the other robots in the formation, and omnidirectional visual servoing for tracking and collision avoidance.

The problem of controlling a formation of ground and aerial vehicles is gaining significant importance in the control and robotics communities, thanks to recent advances in communications and computer vision. Examples of formation control in man-made devices include air-traffic control, satellite clustering, automatic highways, and mobile robotics.

Previous work in formation control assumes that communication among robots is possible and concentrates on aspects of the problem such as stability and controller synthesis. Swaroop et al. [1] proposed the notion of string stability for line formations and derived sufficient conditions for a formation to be string stable. Tanner et al. [2] concentrated on formations in acyclic graphs and proposed the notion of leader-to-formation stability, based on the effect of feedback and feedforward on the input-to-state stability of the formation.

Fax et al. [3] analyzed the stability of formations in arbitrary graphs and proposed a Nyquist-like stability criteria that can be derived from the spectral properties of the graph Laplacian. Stipanovic et al. [4] studied the design of decentralized control laws that result in stable formations, provided that the leader's desired velocity is known.

In the absence of communication, the formation control problem becomes quite challenging from a sensing viewpoint due to the need for simultaneous estimation of the motion of multiple moving objects. Das et al. [5] tackle vision-based formation control with feedback-linearization by employing a clever choice of coordinates in the configuration space. They mitigate sensing difficulties by painting each leader of a different color and then using color tracking to detect and track the leaders. Our work in [6] showed the possibility of using central panoramic cameras to estimate the position and velocities of multiple moving objects based on their optical flows (Figure 2). In more recent work, Cowan et al. [7] proposed two different controllers for image-based leader-follower formation control. One is based on feedback-linearization and the other combines Luenberger observers with a linear controller.



**Special Issue
on Panoramic Robotics**

© DIGITAL STOCK, 1996

In this article, we present a new approach to vision-based formation control of nonholonomic robots equipped with central panoramic cameras in which the detection and tracking of the leaders is based solely on their motions on the image plane (Figure 2). Our approach is to translate the formation control problem from the configuration space into a separate visual servoing control task for each follower. We show how to estimate the position of each leader in the image plane of the follower by using a rank constraint on the central panoramic optical flow across multiple frames. We also derive the leader-follower dynamics in the image plane of each follower for a calibrated central panoramic camera undergoing planar motion. We then show that the closed-loop dynamics after feedback linearization suffer from degenerate configurations due to the nonholonomic constraints of the robots and the nonlinearity of the central panoramic projection model. We, therefore, design a nonlinear tracking controller that avoids such degenerate configurations while maintaining the formation input-to-state stability. That is, we guarantee that the tracking errors are bounded when the leader velocities are. Our control law naturally incorporates collision avoidance by exploiting the geometry of central panoramic cameras. Later, we present simulations and experiments validating our omnidirectional vision-based formation control scheme. Finally, we give directions for future research.

Central Panoramic Formation Dynamics

In this section, we derive the central panoramic optical flow equations for a leader-follower configuration moving in the XY plane and present a multibody motion segmentation algorithm for computing the position of the leaders in the image plane of each follower. We assume that the camera is mounted so that its coordinate system coincides with that of the follower, i.e., the optical center is located at $(X, Y, Z) = 0$ in the follower frame and the optical axis is the Z axis [8].

The Optical Flow of the Central Panoramic Camera

Central panoramic cameras are realizations of omnidirectional vision systems that combine a mirror and a lens and have a unique effective focal point. Building on the results of [9], we show in [6] that the image $(x, y)^T$ of a three-dimensional (3-D) point $q = (X, Y, Z)^T \in \mathbb{R}^3$, obtained by a calibrated central panoramic camera with parameter $\xi \in [0, 1]$, can be modeled as a projection onto the surface

$$z = \mathbf{f}_\xi(x, y) \triangleq \frac{-1 + \xi^2(x^2 + y^2)}{1 + \xi\sqrt{1 + (1 - \xi^2)(x^2 + y^2)}},$$

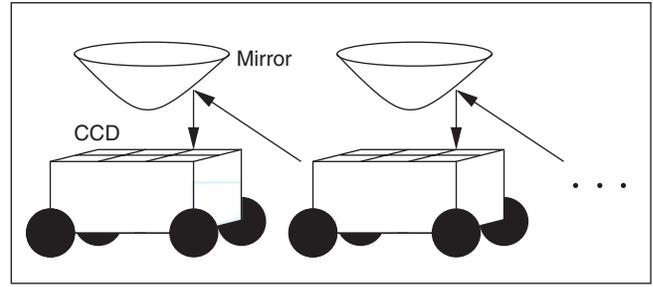


Figure 1. An omnidirectional vision-based formation of mobile robots.

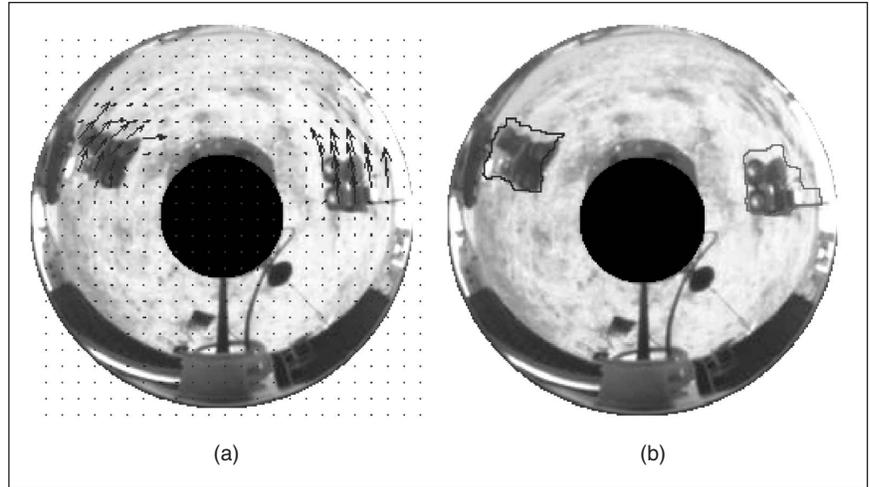


Figure 2. Motion segmentation for two mobile robots based on their omnidirectional optical flows.

followed by orthographic projection onto the XY plane (Figure 3). The composition of these two projections gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{-Z + \xi\sqrt{X^2 + Y^2 + Z^2}} \begin{bmatrix} X \\ Y \end{bmatrix} \triangleq \frac{1}{\lambda} \begin{bmatrix} X \\ Y \end{bmatrix}. \quad (1)$$

Notice that $\xi = 0$ corresponds to perspective projection, while $\xi = 1$ corresponds to paracatadioptric projection (parabolic mirror with orthographic lens).

When the camera moves in the XY plane, its angular and linear velocities are given by $\Omega = (0, 0, \Omega_z)^T \in \mathbb{R}^3$ and $V = (V_x, V_y, 0)^T \in \mathbb{R}^3$, respectively. Relative to the camera, the point q evolves as $\dot{q} = \Omega \times q + V$. This induces a motion in the central panoramic image plane, the so-called “central panoramic optical flow,” which can be computed by differentiating (1) with respect to time. We show in [6] that the optical flow $(\dot{x}, \dot{y})^T$ induced by a central panoramic camera undergoing a planar motion (Ω, V) is given by

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} \Omega_z + \frac{1}{\lambda} \begin{bmatrix} 1 - \rho x^2 & -\rho xy \\ -\rho xy & 1 - \rho y^2 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \end{bmatrix}, \quad (2)$$

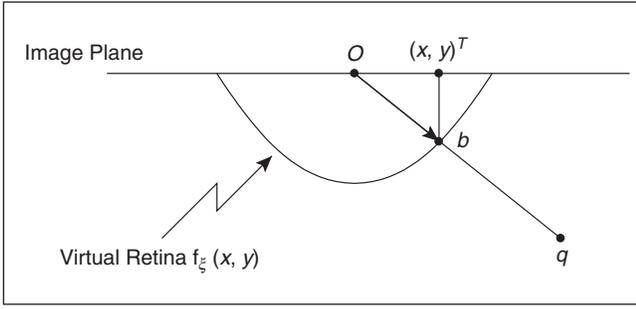


Figure 3. Showing the curved virtual retina in the projection model of central panoramic cameras.

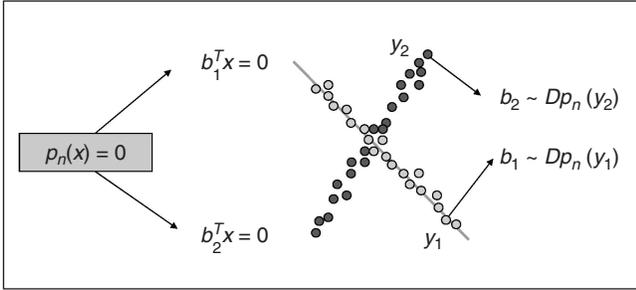


Figure 4. GPCA: segmentation and estimation of multiple subspaces from sampled data by polynomial fitting and differentiation.

where $\lambda = -Z + \xi\sqrt{X^2 + Y^2 + Z^2}$ is an unknown scale factor, $z = \mathbf{f}_\xi(x, y)$ and $\rho \triangleq \xi^2/(1 + z)$.

Central Panoramic Motion Segmentation

Consider a central panoramic camera observing k leaders moving in the XY plane. We now describe how to estimate the image positions of the leaders from measurements of their optical flows across multiple frames. To this end, let $(x_i, y_i)^T$, let $i = 1, \dots, n$, be a pixel in the zeroth frame associated with one of the leaders, and let $(\dot{x}_{ij}, \dot{y}_{ij})^T$ be its optical flow in frame $j = 1, \dots, m$ relative to the zeroth. From (2) we have $[\dot{x}_{ij}, \dot{y}_{ij}] = S_i M_j^T$ where

$$S_i = \begin{bmatrix} x_i & -y_i & \frac{1-\rho_i x_i^2}{\lambda_i} & \frac{-\rho_i x_i y_i}{\lambda_i} & \frac{1-\rho_i y_i^2}{\lambda_i} \end{bmatrix} \in \mathbb{R}^{1 \times 5},$$

$$M_j = \begin{bmatrix} 0 & \Omega_{zj} & V_{xj} & V_{yj} & 0 \\ \Omega_{zj} & 0 & 0 & V_{xj} & V_{yj} \end{bmatrix} \in \mathbb{R}^{2 \times 5}. \quad (3)$$

Let $\dot{x}_{(\cdot)j} = [\dot{x}_{1j}, \dot{x}_{2j}, \dots, \dot{x}_{nj}]^T \in \mathbb{R}^n$ and $\dot{y}_{(\cdot)j} = [\dot{y}_{1j}, \dot{y}_{2j}, \dots, \dot{y}_{nj}]^T \in \mathbb{R}^n$ for $j = 1, \dots, m$. Therefore, the optical flow matrix $W = [\dot{x}_{(\cdot)1} \dot{y}_{(\cdot)1} | \dots | \dot{x}_{(\cdot)m} \dot{y}_{(\cdot)m}] \in \mathbb{R}^{n \times 2m}$ associated with the motion of a single leader satisfies $W = \tilde{S} \tilde{M}^T$, where $\tilde{S} = [S_1^T S_2^T \dots S_n^T]^T \in \mathbb{R}^{n \times 5}$ denotes the structure matrix, and $\tilde{M} = [M_1^T M_2^T \dots M_m^T]^T \in \mathbb{R}^{2m \times 5}$ denotes the motion matrix. We conclude that for a single leader-follower configuration moving in the XY plane, the

collection of central panoramic optical flows across multiple frames lies on a five-dimensional subspace of \mathbb{R}^{2m} .

More generally, the optical flow matrix associated with k independently moving leaders can be decomposed as: $W = \text{diag}(\tilde{S}_1, \dots, \tilde{S}_k) [\tilde{M}_1 \dots \tilde{M}_k]^T = S M^T$, where $S \in \mathbb{R}^{n \times 5k}$ and $M \in \mathbb{R}^{2m \times 5k}$. In practice, however, the optical flow matrix W will not be block diagonal, because the segmentation of the image measurements is not known, i.e., we do not know which pixels correspond to which leader.

Nevertheless, we can still recover the block diagonal structure of W , hence the segmentation of the image measurements, by looking at the structure of W . From our previous discussion, each row of W lives in a subspace of dimension at most five, and so all the rows of W live in k subspaces, one per independently moving object. The problem of estimating and segmenting the motion of these k objects is then equivalent to the problem of estimating and segmenting k subspaces from the rows of W . We solve this problem using an algebraic technique called generalized principal component analysis (GPCA) [11]. The basic idea is that one can fit a polynomial p of degree k to the rows of W , and then obtain a vector normal to each subspace from the derivatives of p (see Figure 4). Since each row of W correspond to one pixel in the image, we segment the image pixels into k groups by assigning each row to the closest subspace.

We use the center of gravity of each group of pixels as the pixel position for that leader. Note that, in practice, there will be an extra group of pixels corresponding to static points in the ground plane, whose motion is simply induced by the motion of the camera. For a formation control scenario with few leaders, we can always identify this group of pixels as the largest one in the image. Since this group does not correspond to a leader, we do not need to compute its center of gravity.

Central Panoramic Leader-Follower Dynamics

Consider now the following nonholonomic ‘‘unicycle’’ model for the dynamics of each leader ℓ and follower f : $\dot{X}_i = v_i \cos \theta_i$, $\dot{Y}_i = v_i \sin \theta_i$, $\dot{\theta}_i = \omega_i$, where $i = \ell, f$, $(X_i, Y_i, \theta_i) \in SE(2)$ is the position and orientation of robot i in the inertial frame, and the inputs v_i and ω_i are the linear and angular velocities, respectively. We showed in [8] that the relative angular and linear velocities of the leader relative to the follower $\Omega_{\ell f} \in \mathbb{R}$ and $V_{\ell f} \in \mathbb{R}^2$, respectively, are given by $\Omega_{\ell f} = \omega_\ell - \omega_f$, and $V_{\ell f} = -[v_f, 0]^T + F_{\ell f}$, where $F_{\ell f} \in \mathbb{R}^2$ is a function of the relative position, orientation, and velocity of the leader in the follower frame.

Consider now a central panoramic camera mounted onboard each follower. Since we are assuming that the camera coordinate system coincides with that of the follower, we can replace the previous expressions for $\Omega_{\ell f}$ and $V_{\ell f}$ in (2) to obtain the optical flow of a pixel associated with leader ℓ in the image plane of follower f . Furthermore, since $z = \mathbf{f}_\xi(x, y)$ and $\lambda = Z/z$, if we assume a ground plane constraint, i.e., if we assume that $Z = Z_{\text{ground}} < 0$ is known, then we can write the equations of motion of a pixel as the drift-free control system:

$$\begin{aligned}
\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} &= - \begin{bmatrix} \frac{1-\rho x^2}{\lambda} & -\gamma \\ \frac{-\rho xy}{\lambda} & x \end{bmatrix} \begin{bmatrix} v_f \\ \omega_f \end{bmatrix} \\
&+ \begin{bmatrix} \frac{1-\rho x^2}{\lambda} & \frac{-\rho xy}{\lambda} & -\gamma \\ \frac{-\rho xy}{\lambda} & \frac{1-\rho y^2}{\lambda} & x \end{bmatrix} \begin{bmatrix} F_{\ell f} \\ \omega_\ell \end{bmatrix} \\
&= H(x, y)u_f + d_{\ell f} \tag{4}
\end{aligned}$$

where $u_f = (v_f, \omega_f)^T \in \mathbb{R}^2$ is the control action for the follower, and $d_{\ell f} \in \mathbb{R}^2$ can be thought of as an external input that depends on the state and control action of the leader and the state of the follower.

Omnidirectional Visual Servoing

In this section, we design a control law u_f for each follower to keep a desired distance r_d and angle α_d from each leader in the image plane; i.e., we assume that we are given a desired pixel location (x_d, y_d) for each leader, where $(x_d, y_d) = (r_d \cos(\alpha_d), r_d \sin(\alpha_d))$.

Visual Servoing by Feedback-Linearization

Let us first apply feedback-linearization to the control system (4) with output $(x, y)^T$. We observe that the system has a well defined vector relative degree of (1, 1) for all pixels (x, y) such that $H(x, y)$ is of rank 2, i.e., whenever $x \neq 0$ and $x^2 + y^2 \neq 1/\xi^2$. In this case, the relative degree of the system is $1 + 1 = 2$, thus the zero dynamics of the system are trivially exponentially minimum phase. Therefore, the control law

$$u_f = -H(x, y)^{-1} \left(d_{\ell f} + \begin{bmatrix} k_1(x - x_d) \\ k_2(y - y_d) \end{bmatrix} \right) \tag{5}$$

results in a locally exponentially stable system around (x_d, y_d) whenever $k_1 > 0$ and $k_2 > 0$.

Notice, however, that (5) is undefined whenever $x = 0$ or $x^2 + y^2 = 1/\xi^2$. The first degenerate configuration $x = 0$ arises from the nonlinearity of the central panoramic projection model and the nonholonomic constraints of the robots. For instance, consider a (static) point in the ground for which $x = 0$. Then, the y component of the flow \dot{y} is zero. Such a flow can be generated by purely translating the follower, by purely rotating the follower, or by an appropriate rotation-translation combination. In other words, given the optical flow of that pixel, we cannot tell whether the follower is rotating or translating. Notice also that, due to the nonholonomic constraints of the robots, if $x = 0$ and $y - y_d \neq 0$, the robot cannot instantaneously compensate the error, since it cannot translate along its Y axis. On the other hand, the second degenerate configuration $x^2 + y^2 = 1/\xi^2$ corresponds to the set of pixels on the outer circle of an omnidirectional image (Figure 2). These pixels are projections of 3-D points

The problem of controlling a formation of ground and aerial vehicles is gaining significant importance in the control and robotics communities, thanks to recent advances in communications and computer vision.

at infinity, i.e., they correspond to the horizon $z = 0$. Therefore, the degenerate configuration $x^2 + y^2 = 1/\xi^2$ is not so critical from a control point of view, because it can be avoided by assuming a finite arena. We therefore assume that $x^2 + y^2 \leq r_{\max}^2 < 1/\xi^2$, from now on.

Visual Servoing by Nonlinear Feedback

Although the control law (5) guarantees locally that $(x(t), y(t))$ converges to (x_d, y_d) asymptotically, this requires that $x(t) \neq 0$ for all t and $x_d \neq 0$. Therefore,

- ♦ one cannot specify a desired formation in which $x_d = 0$, that is a formation in which the leader is directly to the left or to the right of the follower
- ♦ even if $x_d \neq 0$, the controller will saturate when the leader crosses the follower's Y axis at $x = 0$.

Since the latter case is fairly common in most formation configurations, we now design a slightly different controller that avoids this degeneracy, while maintaining the input-to-state stability of the formation. That is, we guarantee that the tracking errors are bounded when the leader velocities are. We first rewrite the leader-follower dynamics in polar coordinates (r, α) , so as to exploit the geometry of the central panoramic camera. The dynamics become $(\dot{r}, \dot{\alpha})^T = \tilde{H}(r, \alpha)u_f + \tilde{d}_{\ell f}$, where $\tilde{H}(r, \alpha)$ and $\tilde{d}_{\ell f}$ are polar coordinate versions of $H(x, y)$ and $d_{\ell f}$, respectively. Rather than exactly inverting the dynamics as in (5), we use the pseudo-feedback linearizing control law:

$$u_f = \begin{bmatrix} \frac{\lambda \cos(\alpha)}{(1-\rho r^2)} & 0 \\ \frac{\sin(\alpha) \cos(\alpha)}{r(1-\rho r^2)} & 1 \end{bmatrix} \left(\begin{bmatrix} k_1(r - r_d) \\ k_2(\alpha - \alpha_d) \end{bmatrix} + \tilde{d}_{\ell f} \right). \tag{6}$$

With this controller, the closed-loop dynamics on the tracking errors $e_r = r - r_d$ and $e_\alpha = \alpha - \alpha_d$ become: $\dot{e}_r = -k_1 \cos^2(\alpha) e_r + \sin^2(\alpha) d_r$, where d_r is the r component of $\tilde{d}_{\ell f}$, and $\dot{e}_\alpha = -k_2 e_\alpha$. Therefore, $\alpha(t)$ converges to α_d asymptotically when $k_2 > 0$. On the other hand, after solving the first order differential equation for the error e_r we obtain $e_r(t) = e_r(t_0) \exp(-k_1 \int_{\tau=t_0}^t \cos^2(\alpha(\tau)) d\tau) + \int_{\tau=t_0}^t \sin^2(\alpha(\tau)) d_r(\tau) \exp(-k_1 \int_{\sigma=\tau}^t \cos^2(\alpha(\sigma)) d\sigma) d\tau$. A simple calculation shows that $|d_r(t)| \leq |v_\ell(t)| / Z + |\omega_\ell(t)|$. Thus, if $k_1 > 0$, $\alpha(t) \neq \pm\pi/2$ from some t on, and the leader velocities (v_ℓ, ω_ℓ) are uniformly bounded, then the formation is leader-to-formation stable (LFS) [2]. Now, since

$\alpha(t) = \alpha_d + e_\alpha(t_0) \exp(-k_2(t - t_0))$, the formation is LFS except when $\alpha_d = \pm\pi/2$ and $e_\alpha(t_0) = 0$.

Notice that the controller (6) is discontinuous at $e_\alpha = \pm\pi$ due to the identification of \mathbb{S}^1 with \mathbb{R} , together with the fact that the seemingly continuous feedback term $k_2 e_\alpha$ does not respect the underlying topology of \mathbb{S}^1 . One could use smooth feedback instead, e.g. $k_2 \sin(\alpha)$, at the cost of a spurious critical point at $\pm\pi$. Since the topology of the annulus dictates that such spurious critical points are inevitable for smooth vector fields, we prefer the discontinuous controller (6) at the benefit of greater performance.

Estimation of the Feedforward Term

In order to implement either controller (5) or controller (6), we need to feedforward the unknown external input $d_{\ell f} \in \mathbb{R}^2$. Although this term is a function of the state and control of the leader and the state of the follower, we do not need to measure any of these quantities. Instead, we only need to estimate the two-dimensional vector $d_{\ell f}$, which can be easily done from the output of the motion segmentation algorithm developed previously. To this end, let (x_w, y_w) and (\dot{x}_w, \dot{y}_w) be the position and optical flow of a pixel that corresponds to a static 3-D point in the world such that $x_w \neq 0$. Similarly, let (x_ℓ, y_ℓ) and $(\dot{x}_\ell, \dot{y}_\ell)$ be the position and optical flow of a pixel that corresponds to a 3-D point on leader ℓ . From (4), where the second term is zero because the point in 3-D space is static (i.e. $(v_\ell, \omega_\ell) = (0, 0)$), the external disturbance can be estimated as: $d_{\ell f} = [\dot{x}_\ell \ \dot{y}_\ell]^T - H(x_\ell, y_\ell)H(x_w, y_w)^{-1}[\dot{x}_w \ \dot{y}_w]^T$.

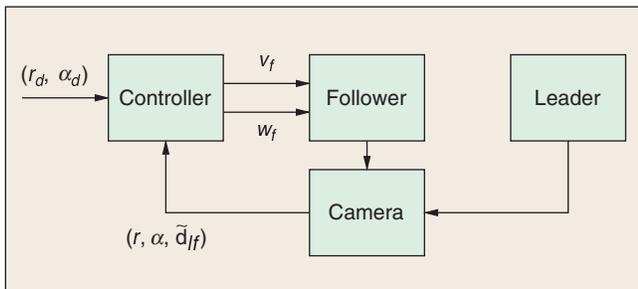


Figure 5. Omnidirectional vision-based formation control scheme.

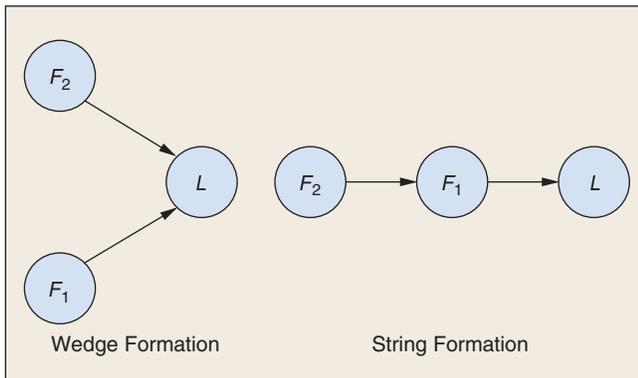


Figure 6. Formation configurations.

Notice that in the presence of noise, one may improve the estimation of the feedforward term by using more than one pixel corresponding to the static background and solving the previous equation in a least squares sense.

Collision Avoidance

Although the control law (5) guarantees local stability of the leader-follower formation, it does not guarantee that the follower will not run into the leader. For example, imagine that the follower is initially in front of the leader and that the desired formation is with the follower behind the leader. Since the closed-loop dynamics are linear in the error $(x - x_d, y - y_d)$, the follower will apply a negative linear speed, and will most likely run into the leader.

Thanks to the geometry of central panoramic cameras, collisions can be avoided by ensuring that the leader stays far enough away from the center of the image. Effectively, our choice of image coordinates (r, α) for the controller (6) reveals the safe configurations as a simple constraint on r , namely $r_{\min} \leq r \leq r_{\max}$. Furthermore, the control law (6) is the gradient of a potential function $V(r, \alpha) = \frac{1}{2}(k_1(r - r_d)^2 + k_2(\alpha - \alpha_d)^2)$, which points transversely away from the safety boundary, and has a unique minimum at (r_d, α_d) (assuming $r_d > r_{\min}$). Therefore, the control law (6), or a polar coordinate version of controller (5), naturally incorporate collision avoidance by exploiting the geometry of central panoramic cameras. In order to further guarantee collision avoidance, one could modify the potential function $V(r, \alpha)$ to yield a proper navigation function, as suggested in Cowan et al. [10], and then use the gradient for the modified V as the control law.

Experimental Results

We evaluated the performance of the proposed multi-body motion segmentation algorithm in the case where two independently moving mobile robots are viewed by a static camera. We grabbed 18 images of size 240×240 pixels from the omnidirectional camera at a framerate of 5 Hz. The optical flow was computed using Black's algorithm, available at <http://www.cs.brown.edu/people/black/ignc.html>.

Figure 2 shows a sample of the motion segmentation based on the optical flow. The optical flow generated by the two moving robots is shown in (a), and the segmentation of the pixels corresponding to the independent motions is in (b). The two independently moving robots are segmented very well from the static background.

We also tested our omnidirectional vision-based formation control scheme (Figure 5) by having three nonholonomic robots start in a Wedge Formation and then follow a String Formation with $(r_d, \alpha_d) = (1/\sqrt{2}, 0)$ (Figure 6). Since $\alpha_d = 0$, we choose to use controller (5) in polar coordinates with the parameters $\xi = 1$, $k_1 = 2.5$ and $k_2 = 1.76$. Figure 7 shows the simulation results. For $t \in [0, 29]$, the leader moves with $v_\ell = 0.5$ and $\omega_\ell = 0$, and the followers move from the initial configuration to the desired one. Notice how the followers automatically avoid collision when Follower1 tries to move in between Follower2 and the leader. For $t \in [29, 36]$,

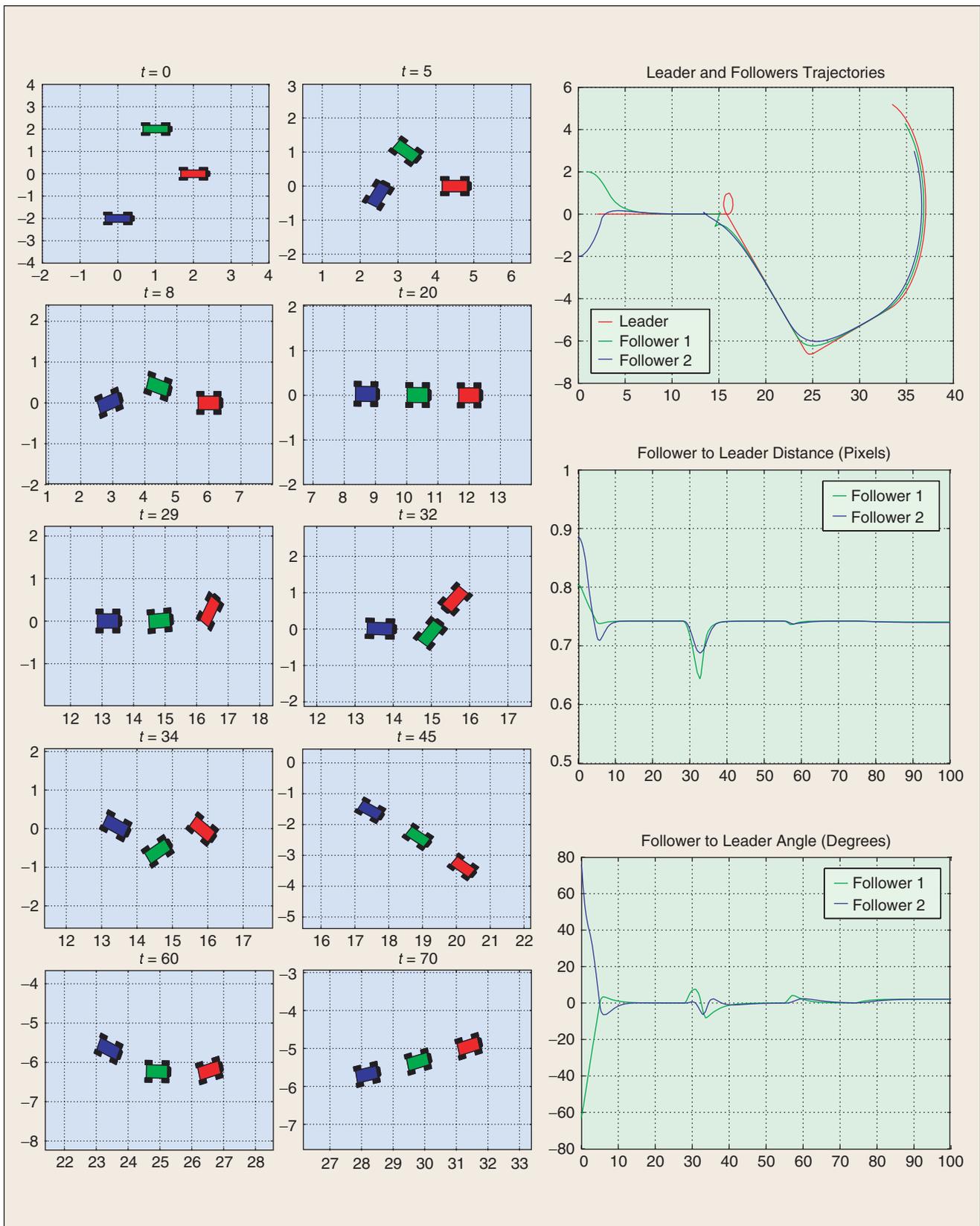


Figure 7. Simulation results for a string formation. For $t \in [0, 10]$ the followers move from their initial wedge formation to the desired spring formation, while avoiding a collision when Follower1 moves in between Follower2 and the leader. The leader abruptly rotates for $t \in [29, 36]$, but the followers are able to both avoid collision and later return to the desired line. For $t > 36$, they maintain their formation into a line, circle, line and a circle. Notice that it is not possible to maintain zero angular error during circular motion, because of the nonholonomic kinematic constraints.

the leader changes its angular velocity to $w_\ell = 1$, thus moving in a circle. Follower1 starts rotating to the right to follow the leader, but soon realizes that the leader is coming towards it and backs up to avoid collision. For $t \in [36, 55]$, the leader changes its angular velocity to $w_\ell = 0$, and the followers are able to return to the desired formation. For $t \in [55, 60]$, the leader turns at $w_\ell = 0.5$ and the followers are able to keep the formation. For $t \in [60, 100]$, the leader turns at $w_\ell = 0.1$ and the followers maintain the formation into a line and a circle.

Conclusions and Future Work

We have presented a new approach to formation control of nonholonomic mobile robots equipped with central panoramic cameras. Our approach uses motion segmentation techniques to estimate the position of each leader and omnidirectional visual servoing for tracking and collision avoidance. We showed that direct feedback-linearization of the leader-follower dynamics leads to asymptotic tracking but suffers from degenerate configurations. We therefore presented a nonlinear controller that avoids such singularities but can only guarantee input-to-state stability of the formation.

Future work will include combining the two controllers presented in this article in a hybrid theoretic formulation that allows the design of a feedback control law that avoids singularities and guarantees asymptotic tracking. We would also like to explore the design of alternative control laws that do not use optical flow estimates in the computation of the feedforward term. We also plan to implement our formation control scheme on the Berkeley test bed of unmanned ground and aerial vehicles.

Acknowledgments

We thank Dr. Noah Cowan for his insightful comments on the preparation of the final manuscript. We also thank the support of ONR grant N00014-00-1-0621 and DARPA grant F33615-98-C-3614.

References

- [1] D. Swaroop and J. Hedrick, "String stability of interconnected systems," *IEEE Trans. Automat. Control*, vol. 41, no. 3, pp. 349–357, 1996.
- [2] H. Tanner, G. Pappas, and V. Kumar, "Leader-to-formation stability," *IEEE Trans. Robot. Automat.*, vol. 20, no. 3, pp. 443–455, 2004.
- [3] A. Fax and R. Murray, "Graph Laplacians and stabilization of vehicle formations," in *Proc. Int. Federation Automatic Control World Congr.*, 2002.
- [4] D. Stipanovic, G. Inalhan, R. Teo, and C. Tomlin, "Decentralized overlapping control of a formation of unmanned aerial vehicles," in *Proc. IEEE Conf. Decision Control*, 2002, pp. 2829–2835.
- [5] A. Das, R. Fierro, V. Kumar, J. Ostrowski, J. Spletzer, and C. Taylor, "A vision-based formation control framework," *IEEE Trans. Robot. Automat.*, vol. 18, no. 5, pp. 813–825, 2002.
- [6] O. Shakernia, R. Vidal, and S. Sastry, "Multibody motion estimation and segmentation from multiple central panoramic views," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2003, vol. 1, pp. 571–576.
- [7] N. Cowan, O. Shakernia, R. Vidal, and S. Sastry, "Vision-based follow-the-leader," Submitted to: *IEEE Int. Conf. Intelligent Robotics and Systems (IROS)*, 2003, vol. 2, pp. 1796–1801.
- [8] R. Vidal, O. Shakernia, and S. Sastry, "Formation control of nonholonomic mobile robots with omnidirectional visual servoing and motion segmentation," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2003, vol. 1, pp. 584–589.
- [9] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *European Conf. Computer Vision*, 2000, pp. 445–461.
- [10] N. Cowan, J. Weingarten, and D. Koditschek, "Visual servoing via navigation functions," *IEEE Trans. Robot. Automat.*, vol. 18, no. 4, pp. 521–533, 2002.
- [11] R. Vidal, Y. Ma, and J. Piazzi, "A new GPCA algorithm for clustering subspaces by fitting, differentiating, and dividing polynomials," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2004, vol. 1, pp. 510–517.

René Vidal received his Ph.D. degree in electrical engineering and computer sciences from UC Berkeley in 2003. He is currently an assistant professor of biomedical engineering and a member of the Center for Imaging Science at the Johns Hopkins University. His research interests include GPCA, segmentation of dynamic scenes, multiple view geometry, omnidirectional vision, vision-based control of mobile robots, hybrid systems identification, and game theory. Dr. Vidal is the recipient of the Best Paper Award Honorable Mention at ECCV 2004, the 2004 Sakrison Memorial Prize, and the 2003 Eli Jury Award. He is a student member of IEEE.

Omid Shakernia received his Ph.D. degree in electrical engineering and computer sciences from UC Berkeley in 2003. His principle research interest at Berkeley was computer vision based coordination and control of unmanned vehicles. He is currently at Northrop Grumman Corporation, developing autonomy algorithms for unmanned aerial vehicles that incorporate computer vision and flight control technologies.

Shankar Sastry received his Ph.D. degree in 1981 from UC Berkeley. He was on the faculty of MIT as an assistant professor from 1980–1982, Harvard University as a chaired Gordon McKay professor in 1994, and is currently NEC professor and chair of EECS at UC Berkeley. He has coauthored over 250 technical papers and eight books, including *Nonlinear Systems: Analysis, Stability and Control*, published by Springer-Verlag in 1999. His areas of research are embedded and autonomous software, computer vision, nonlinear and adaptive control, robotic telesurgery, control of hybrid systems, embedded systems, sensor networks, and biological motor control. Dr. Sastry was elected into the National Academy of Engineering in 2001.

Address for Correspondence: René Vidal, Center for Imaging Science, Johns Hopkins University, 308B Clark Hall, 3400 N. Charles St., Baltimore, MD 21218 USA. E-mail: rvidal@cis.jhu.edu.