

IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'2002
 pp. 390-395, Lausanne, Switzerland, October 2002

Choice of image features for depth-axis control in image based visual servo control

Robert Mahony[♣], Peter Corke[♡] and Francois Chaumette[♠]

[♣] *Dep. Engineering,
 Australian National University
 Acton, ACT
 Australia.
 Robert.Mahony@anu.edu.au*

[♡] *CSIRO,
 Manufacturing Science & Technology
 Pullenvale, Queensland.
 Australia.
 peter.corke@csiro.au*

[♠] *IRISA
 INRIA
 Campus de Beaulieu, Rennes
 France.
 Francois.Chaumette@irisa.fr*

Abstract

This paper is concerned with choosing image features for image based visual servo control and how this choice influences the closed-loop dynamics of the system. In prior work, image features tend to be chosen on the basis of image processing simplicity and noise sensitivity. In this paper we show that the choice of feature directly influences the closed-loop dynamics in task-space. We focus on the depth axis control of a visual servo system and compare analytically various approaches that have been reported recently in the literature. The theoretical predictions are verified by experiment.

1 Introduction

Image-based visual servo (IBVS) control is a technique whereby features f derived solely from image plane information (image features) are servo controlled to a desired goal configuration [11]. Tasks in Cartesian space are re-posed as a servo control task directly in image feature space. The classical IBVS approach is based on controlling the image feature kinematics. These are given by the expression

$$\dot{f} = J \begin{pmatrix} V \\ \Omega \end{pmatrix} \quad (1)$$

where f is the vector of image features, (V, Ω) is the camera's velocity screw and the matrix J is the feature Jacobian or interaction matrix and is a function both of the image features and of the unknown Cartesian pose of the camera. A linearising control in image feature space

$$\begin{pmatrix} V \\ \Omega \end{pmatrix} = J^{-\dagger}(f - f^*), \quad (2)$$

where $J^{-\dagger}$ denotes the pseudo-inverse of J , is the typical control strategy used. It is a known feature of IBVS that the closed-loop system may display undesirable dynamics in task-space even though the dynamics in the image space are linear [3].

To date, most work in image-based visual servo control has made use of point features [5,6,9,14], typically the coordinates of centroids of markers or corners. Other features have been proposed in the literature including the distance between two points in the image plane and the orientation of the line connecting those two points [6], perceived edge length [17], the relative areas of two projected surfaces [17], the centroid and higher order moments of a projected surface [1,17,18], or the parameters of lines or ellipses in the image plane [2,5]. More abstract features such as rotation and depth information derived from homography decomposition have also been demonstrated [13]. In some cases specific features such as target surface area [4] or the norm of a weighted centroid feature [8] have been used in order to better control the depth axis of the closed-loop system. From an image processing point of view the use of a feature obtained by an integration process, such as a moment, is likely to be more robust than a measure obtained by a differentiation process such as a point or line. A key observation is that recent methods tend to use an amalgamation of various image features, such as points, angles, surface area etc., to servo control the system rather than relying on a single class of features [4]. Despite the wide range of features considered, to the authors knowledge, there has been no explicit work that analyzes in detail the impact of the choice of image feature on the closed-loop system dynamics in a task space.

In designing a visual servo control system it is appropriate to ask the question: "What is a good set of image features to select for IBVS control?". In practice, it is possible to assign a given dynamic response to the image feature kinematics, however, it is the task kinematics of the robot in Cartesian space that are of most interest in a performance analysis. Ideally, one would wish that the dynamics in the image feature space correspond to equivalent dynamics in the task-space. In this paper, we analyse a specific class of IBVS problems, that in which the camera is

constrained to move in the depth or z -axis. Such motion control is always problematic in IBVS schemes since changes in z lead to relatively small and highly coupled changes in image features. We show that it is possible to find image features for which a linear control design in image feature space does indeed lead to asymptotically stable linear dynamics of the closed-loop system in task-space. A number of different image features and their associated dynamics are analyzed in detail to provide an overview of the problem and cover most of the features recently proposed in the literature. The theoretical predictions are verified by experiment.

The next section more formally articulates the problem we are seeking to address. Section 3 treats a number of possible image features analytically to determine their functional relationship to depth and derives analytic expressions for the closed-loop system response of an IBVS control design. The candidate image features fall into distinct families. Section 4 presents an experimental study that validates the theoretical predictions of Section 3 for an example feature from each family considered. Finally Section 5 summarizes our findings and discusses future work.

2 Formulation of problem

In order to obtain a tractable problem we restrict our attention to the case where only motion along the optical (also termed depth or z) axis of the camera is allowed.

To motivate the results presented later in the paper we begin by recalling a particular case study of classical IBVS control [3]. Consider a camera located directly above a square planar target (of side length 1 for simplicity) perpendicular to the optical axis of the camera. To simplify the analysis we assume that the focal length of the camera is unity. The pixel coordinates of the image points are $(X_i, Y_i) = (\pm a, \pm a)$ where $a = 1/(2z)$ and z denotes the depth of the target from the camera. Choose a visual feature vector $f = (X_1, \dots, X_4, Y_1, \dots, Y_4)$ then the image feature kinematics are given by

$$\dot{f} = J_f \begin{pmatrix} V \\ \Omega \end{pmatrix}$$

where J_f is the image feature Jacobian or interaction matrix. An analytic expression for the interaction matrix is given by Chaumette [2]. Let z_0 denote the initial depth and $f_0 = f(z_0)$ denote the initial image feature. Choose the target feature associated with motion of the camera to a desired depth z^* and rotation of the camera through π radians. Due to the particular geometry of the example it is easily verified that $f^* = -\frac{z_0}{z^*} f_0$. The image error is defined to be

$$e = f - f^*.$$

In classical image based visual servo the image feature error kinematics are assigned stable dynamics. Typically, a linearizing control is used

$$\begin{pmatrix} \dot{V} \\ \dot{\Omega} \end{pmatrix} = -\beta J_f^\dagger e \quad (3)$$

where $J_f^{-\dagger}$ denotes the pseudo-inverse of J_f and $\beta > 0$ is a positive constant. This control assigns the kinematics $\dot{e} = -\beta e$ to the image feature error. An analytic form for J_f^\dagger is given by Chaumette [2]. From this information it is easily verified that the IBVS linearising control design leads to a control input $V_z = \beta(z + \frac{z^2}{z^*})$. The z -axis dynamics in Cartesian task space are

$$\dot{z} = \beta(z + \frac{z^2}{z^*}). \quad (4)$$

We have shown that the linear dynamics assigned in image feature space lead to an unstable non-linear quadratic ODE in task-space.

The above example is an extreme example of the well known problem with classical IBVS — the camera trajectory is not optimal in Cartesian space. For reasons of simplicity and performance, it is desirable to assign linear task-space dynamics

$$\dot{z} = -\gamma(z - z^*), \quad (5)$$

for $\gamma > 0$ some positive constant. At the same time it is desirable to preserve the linear dynamics in image feature space since this significantly simplifies the control design. Thus, we ask the following question:

Is it possible to find image features such that assigning asymptotically stable linear dynamics in image feature space leads to asymptotically stable linear task dynamics in Cartesian space?

Consider the specific case where only motion in the z -axis is possible. The linearizing control for the image feature error is given by

$$V_z = -\beta J_z^{-1} e, \quad \text{where } J_z = \frac{\partial e}{\partial z}, \quad (6)$$

for a constant $\beta > 0$. Equating the desired linear task dynamics Eq. 5 with the linearizing control design Eq. 6 one obtains

$$\frac{1}{e} \frac{\partial e}{\partial z} = \frac{\beta}{\gamma(z - z^*)}$$

As a consequence one has

$$e = k_0(z - z^*)^{(\beta/\gamma)} \quad (7)$$

where k_0 is some constant. We have shown that, (in the case where only motion in the z -axis is possible.)

if linear dynamics in both the task and the image feature error are desired, it is sufficient to choose the image feature error of the form Eq. 7 with $\beta/\gamma > 0$. In practice, it may be difficult to find an image feature with the exact scaling features of Eq. 7. However, the simplest case, where the linear rate of convergence in both the image feature and the task-space are equivalent $\beta = \gamma$, is of considerable interest. In this case one has

$$e \equiv f - f^* = k_0(z - z^*), \quad J_z \equiv k_0.$$

This has a particularly nice form in that the interaction matrix has constant sensitivity with depth, a property that improves the numerical robustness of the IBVS control law obtained. It is important to note that using an image feature of this form *is not* equivalent to 3D position based visual servo control. The constant k_0 or the power β/γ need not be known and the error e is constructed explicitly from image data. Of course, it is necessary to use an estimate of the image Jacobian or interaction matrix in order to compute the control law and this usually depends in part on the parameters of the image feature as well as the unknown depth z . This is a common feature of all IBVS algorithms and a number of methods exist to overcome the difficulties [10, 12, 15, 16].

3 The dynamics of the optical axis in closed-loop IBVS

In this section a comparative theoretical analysis of a range of characteristic image feature errors that have been proposed in recent literature is undertaken. We focus on understanding the dynamic response of the closed-loop IBVS control system in the ideal case where the exact linearizing control law can be computed.

3.1 Image feature errors proportional to z^{-1}

Classical IBVS control based on point features leads to an error criteria along the optical axis that is inversely proportional to the average depth of the target. This follows from the dependence of the depth sensitivity on distance between points in the image plane. This is a scaling property of any line segment or linear measure derived from the image for a perspective projection camera subject to movement in the z -axis. Other examples of image features with the same scaling property are; direct measurement of the length of a line segment $e = l - l^*$, or the radius of the smallest disk in the image plane that contains the image. A robust measure that can be used is the square root of the zero order moment [4]

$$e = \sqrt{m_{00}} - \sqrt{m_{00}^*}. \quad (8)$$

For all these cases one has

$$e \approx \frac{k}{z} - A$$

where A is a constant linked to the goal image feature while $k > 0$ is a proportionality constant that depends on the nature of the target image and camera geometry. Computing the z dynamics for a linearizing control in image feature space one obtains task-space dynamics

$$\dot{z} = z - \frac{A}{k}z^2$$

in the optical axis of the camera frame. Note the equivalence of this with Eq. 4 derived directly from the full image Jacobian structure. The solution of this ODE is easily computed using partial fraction techniques

$$z(t) = \frac{k}{A + ke^{(c-t)}}$$

where c is a constant depending on the initial conditions of the system.

The solutions of this system display two behaviours: Firstly, when $A \geq 0$ or $A < -ke^c$ then $z(t) \rightarrow k/A$ and the asymptotic convergence is linear

$$\left(z(t) - \frac{k}{A}\right) \propto e^{-t}, \quad t \rightarrow \infty.$$

Alternatively, if $-e^c < A < 0$ then $z(t) \rightarrow \infty$ for $t \rightarrow c - \ln(-1/A)$. This is the explicit form of the finite-time escape behaviour that was observed in the Chaumette conundrum [3, 4].

3.2 Image feature errors proportional to z^{-2}

A number of authors have considered visual servo control using the surface area of the target directly. The image feature error in this case is

$$e = m_{00} - m_{00}^*. \quad (9)$$

Any image feature that depends on a measure of projected area in the image plane will scale as $1/z^2$ with motion in the optical axis of the camera. Consider an image feature error of the form

$$e \approx \frac{k}{z^2} - A$$

where A and k are constants analogous to those in Section 3.1. Computing the z dynamics $\dot{z} = V_z$ for a linearizing control law in the image feature kinematics leads to task-space dynamics

$$\dot{z} = z - \frac{A}{2k}z^3$$

The solution of this ODE may be computed using partial fraction techniques

$$z(t) = \sqrt{\frac{k}{A + e^{(c-t)}}$$

where c is a constant depending on the initial conditions of the system. Note that the solution obtained is simply the square root of the solution obtained in Subsection 3.1. Clearly, the solution will display the same qualitative behaviour, however, the asymptotic convergence for $A \geq 0$ or $A < -e^c$ is linear with rate $1/2$

$$\left(z(t) - \sqrt{\frac{k}{A}}\right) \propto e^{-t/2}, \quad t \rightarrow \infty.$$

If $-e^c < A < 0$ then $z(t) \rightarrow \infty$ for $t \rightarrow c - \ln(-1/A)$.

3.3 Image feature errors proportional to z^2

It is of interest to consider the inverse of the zero order image moment since it intuitively scales in the same manner as depth (increases for increasing depth and vice versa). The depth control of the scheme recently proposed by Hamel *et al.* [8] is closely related to this form. A general image error of the form

$$e = \frac{1}{m_{00}} - \frac{1}{m_{00}^*} \quad (10)$$

or

$$e = \frac{m_{00}^*}{m_{00}} - 1$$

are examples of such an image feature. A general form for this class of image feature error is

$$e \approx kz^2 - A$$

where A and k are constants analogous to those in above. Computing the z dynamics $\dot{z} = V_z$ for a linearising control law in the image feature kinematics leads to task-space dynamics

$$\dot{z} = \frac{1}{2k} \left(\frac{A}{z} - kz \right)$$

The solution of this is easily computed using a change of variables $\hat{z} = (z^2 - A/k)$ and yields

$$z(t) = \sqrt{\frac{A + e^{(c-t)}}{k}} \quad (11)$$

where c is a constant depending on the initial conditions of the system. The solution obtained is simply the inverse of the case in Subsection 3.2. However, the qualitative behaviour of the solution is considerably improved since the potential singularity is removed. Thus, for all constants A , c and k , $z \rightarrow \sqrt{A/k}$ and

$$\left(z(t) - \sqrt{\frac{A}{k}}\right) \propto e^{-t/2}, \quad t \rightarrow \infty.$$

Remark 3.1 In Section 2 it was observed that for any image feature error of the form $e \propto (z - z^*)^{\beta/\gamma}$ for $\beta/\gamma > 0$ both the image feature kinematics and the task kinematics are linear. In practice, (except in the case discussed in Subsection 3.4) it is difficult to find an image feature error with exactly the desired property and the best possible is something of the form

$$e \propto k_0 z^{\beta/\gamma} - k_0 (z^*)^{\beta/\gamma}$$

analogous to the case (for $\beta/\gamma = 2$) considered in this subsection. It is interesting to note that the closed-loop response (cf. Eq. 11) is close to linear and the asymptotic convergence in task-space is a factor of half the asymptotic convergence in image feature space, as predicted by the development in Section 2.

3.4 Image feature errors proportional to z

It is possible to find image features proportional to the depth parameter z . A useful and robust example is the inverse square root of the zero order moment

$$e = \frac{1}{\sqrt{m_{00}}} - \frac{1}{\sqrt{m_{00}^*}}. \quad (12)$$

Alternatively one could invert any linear measure derived from the image $e = \frac{1}{l} - \frac{1}{l^*}$.

Remark 3.2 Note that the control design is not equivalent to position based visual servo (PBVS) control where the depth is reconstructed and controlled directly. The approach is a true image based visual servo control design, with the additional advantage that the scaling property of the image feature chosen provides closed-loop performance comparable to that of a PBVS system with well calibrated cameras.

The image feature error may be written

$$e \approx kz - A$$

where A and k are constants analogous to those in above. Computing the z dynamics $\dot{z} = V_z$ for a linearising control law in the image feature kinematics leads to task-space dynamics

$$\dot{z} = \frac{1}{k} (A - kz)$$

The solution of this linear ODE is

$$z(t) = \frac{A + e^{(c-t)}}{k}$$

where c is a constant depending on the initial conditions of the system. For all constants A , c and k , $z \rightarrow A/k$ and

$$(z(t) - A/k) \propto e^{-t}, \quad t \rightarrow \infty.$$

The rate of convergence in image space and feature space is equivalent. This is an example of an image feature that satisfies Eq. 7.

3.5 Image feature errors with log dependence z

In all the above cases the image features considered have been proportional to some power of the depth parameter. There has been some recent consideration of using a log ratio of image features [7] in order that polynomial dependence is converted into a scaling factor. For example, consider image feature errors of the form

$$\begin{aligned} e &= \ln\left(\frac{l^*}{l}\right) = \ln\left(\frac{k_1 z}{A_1}\right), \\ e &= \ln\left(\frac{m_{00}}{m_{00}^*}\right) = -2 \ln\left(\frac{k_2 z}{A_2}\right), \\ e &= \ln\left(\frac{\sqrt{m_{00}}}{\sqrt{m_{00}^*}}\right) = -\ln\left(\frac{k_3 z}{A_3}\right), \end{aligned} \quad (13)$$

for suitable constants k_i , A_i . All such image feature errors differ only by a scalar factor. Thus, we consider a class of image feature errors

$$e = \alpha \ln\left(\frac{kz}{A}\right) \quad (14)$$

where A and k are constants analogous to those in above and α is a constant that is derived from the particular feature used. Computing the z dynamics $\dot{z} = V_z$ for a linearizing control law in the image feature kinematics leads to task kinematics

$$\dot{z} = -z \ln\left(\frac{kz}{A}\right).$$

Note that the task kinematics do not depend on the scale factor. This non-linear ODE may be solved analytically using the substitution $x = \ln(kz/A)$

$$z(t) = \frac{A}{k} e^{e^{(c-t)}}$$

where c is a constant depending on the initial conditions of the system. Clearly the system response is non-linear in task-space, however, the qualitative behaviour is good. For all constants A , c and k , $z \rightarrow A/k$ and

$$(z(t) - A/k) \propto \left(\left(e^{e^{-t}} \right)^{e^{ec}} - 1 \right), \quad t \rightarrow \infty.$$

4 Experimental study

In this section an experimental study is presented that verifies the theoretical prediction made in Section 3. Six image features were considered; a classical IBVS point based image feature error, and then one example from each class of image feature errors discussed in Section 3. To simplify the experiment the image features used were based on a calculation of the first order moment of the image surface

Eqn's (8), (9), (10), (12), and (13). A square target orthogonal to the camera's optical axis was used. Figure 1 shows the set point images from the upper limit ($z = 0.7m$) and the lower limit ($z = 0.2m$). For each error criterion the closed-loop response for servo control, firstly from the upper to the lower set points, and then from the lower to upper set points, was measured. In order to compute the exact linearising control a depth estimate is required and this was provided using a classical pose algorithm.

Figures 2 and 3 show the z -axis velocity demand for the two sets of experiments. The control gain β was used to normalise the initial control demand such that all controllers demand the same initial velocity. This helps in relative comparison of the closed-loop response. As expected, the results for the classical point based IBVS error feature and the feature given by (8) give identical results (red and green plots are superposed). Furthermore, comparing Figure 2 and 3, it is seen that the feature given by (12) gives equivalent exponential control in both upward and downward servo control. All the other features swap between being more aggressive and less aggressive depending on whether upward or downward servo control is considered. Thus features which have a fast time-to-convergence for forward motion have a slow time-to-convergence for backward motion, and vice-versa. In addition, in the case of backward motion the error features (8), (9) and (13) lead to an undesirable increase in the control demand during the transient. Figure 4 plots the closed-loop response in Cartesian space against the response in feature space for the upward servo task. The light blue line corresponds to the error feature (12). Its linear form demonstrates the matched exponential convergence of the Cartesian motion and the image feature motion. The curved nature of the other response lines show the non-linear nature of the Cartesian dynamics of all other schemes. As far as noise sensitivity is concerned, we have not been able to find any significant difference between the different features.

5 Conclusion

This paper has shown analytically and by experiment that the choice of image feature has a direct effect on the closed-loop dynamics of the system. For good closed-loop behaviour in the z -axis for IBVS control schemes, one should choose image features that scale as $e \propto z$.

References

- [1] R. Anderson, *A robot ping-pong player. Experiment in real time intelligent control*, MIT Press, Cambridge, 1988.
- [2] F. Chaumette, *La relation vision-commande : théorie et applications à des tâches robotiques*, PhD thesis, Université de Rennes 1, IRISA, July 1990.

[3] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing", In *The Confluence of Vision and Control*, vol. 237 of *LNCIS Series*, pp. 66-78, 1998.

[4] P. Corke, S. Hutchinson: "A new partitioned approach to image-based visual servo control", *IEEE Trans. on Robotics and Automation*, 17(4):507-515, Aug. 2001.

[5] B. Espiau, F. Chaumette, P. Rives, "A new approach to visual servoing in robotics", *IEEE Trans. on Robotics and Automation*, 8(6):313-326, June 1992.

[6] J. Feddema, O. Mitchell, "Vision-guided servoing with feature-based trajectory generation", *IEEE Trans. on Robotics and Automation*, 5:691-700, Oct. 1989.

[7] N. Gans, P. Corke, S. Hutchinson, "Comparison of robustness and performance of partitioned image based visual servo systems", *Proc. Aust. Conf. on Robotics and Automation*, pp. 73-78, Sydney, Nov. 2001.

[8] T. Hamel, R. Mahony, "Visual servoing of an under-actuated dynamic rigid body system: an image based approach", *IEEE Trans. on Robotics and Automation*, April 2002.

[9] K. Hashimoto, T. Kimoto, T. Ebine, H. Kimura, "Manipulator control with image-based visual servo", *Proc. ICRA '91*, pp. 2267-2272, Sacramento, April 1991.

[10] K. Hosoda, M. Asada, "Versatile visual servoing without knowledge of true jacobian", *Proc. IROS '94*, Munchen, Germany, Sept. 1994.

[11] S. Hutchinson, G. Hager, P. Corke, "A tutorial on visual servo control", *IEEE Trans. on Robotics and Automation*, 12(5):651-670, Oct. 1996.

[12] M. Jägersand, O. Fuentes, R. Nelson, "Experimental evaluation of uncalibrated visual servoing for precision manipulation", *Proc ICRA '97*, Vol. 3, pp. 2874-2880, Albuquerque, April 1997.

[13] E. Malis, F. Chaumette, S. Boudet, "2 1/2 D visual servoing", *IEEE Trans. on Robotics and Automation*, 15(2):238-250, April 1999.

[14] N. Papanikolopoulos, P. Khosla, T. Kanade, "Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision", *IEEE Trans. on Robotics and Automation*, 9(1):14-35, Feb 1993.

[15] N. Papanikolopoulos, P. Khosla, "Adaptive robot visual tracking: theory and experiments", *IEEE Trans. on Automatic Control*, 38(3):429-445, 1993.

[16] J. Piepmeier, G. McMurray, H. Lipkin, "A dynamic quasi-Newton method for uncalibrated visual servoing", *Proc. ICRA '99*, pp. 1595-1600, 1999.

[17] L. Weiss, A. Sanderson, C. Neuman, "Dynamic sensor-based control of robots with visual feedback", *IEEE Journal of Robotics and Automation*, 3(5):404-417, Oct. 1987.

[18] B. Yoshimi, P. Allen, "Active uncalibrated visual servoing", *Proc. ICRA '94*, pp. 156-161, San Diego, May 1994.

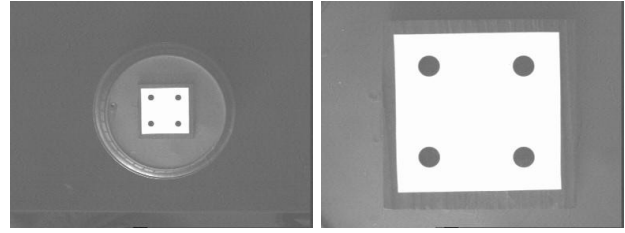


Figure 1: Images for $z=0.7$ m and $z=0.2$ m

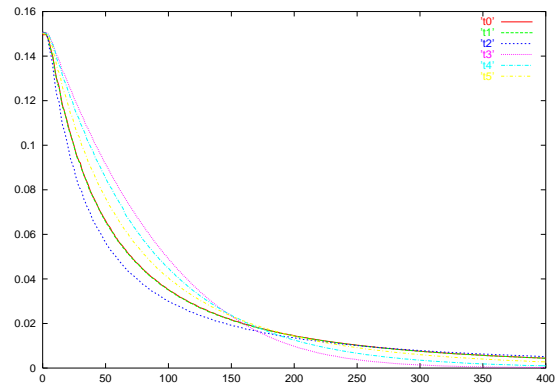


Figure 2: z -axis translation demand versus time to reach $z=0.2$ m from $z=0.7$ m

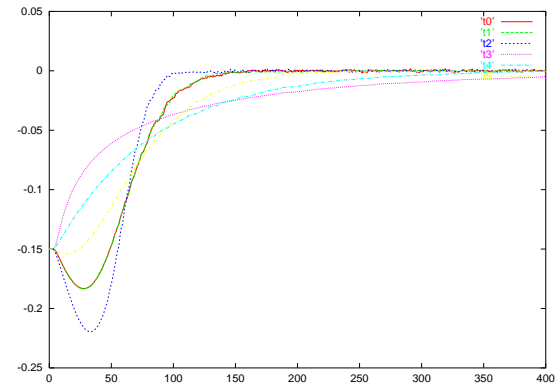


Figure 3: z -axis translation demand versus time to reach $z=0.7$ m from $z=0.2$ m

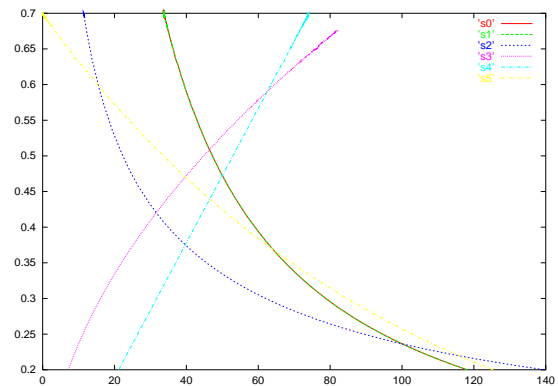


Figure 4: Behaviour of Cartesian target depth versus image feature error