View Variation of Point-set and Line-segment Features

J. Brian Burns, R. Weiss and E.M. Riseman *

COINS Technical Report 90-84

September 7, 1990

Abstract

The recognition of 3D objects becomes much more difficult as the position of the viewer relative to the object becomes less constrained. Part of this difficulty comes from the fact that many important features of an object's image vary with view, and for image features to be effective in the discrimination of 3D objects, the distribution of the feature values over permissible views should tend to be narrow. This paper is a study of the variation, with respect to view, of features defined for projected point sets and line segments. It is first established that general-case view-invariants do not exist for any number of points, given true perspective, weak perspective or orthographic projection models.

The remainder of the paper focuses on feature variation under weak perspective. This approximation to perspective projection is applied extensively in object recognition research to simplify the analysis and computation for 3D object recognition; it produces reasonable results when the camera distance is great enough, relative to the object's extent in depth. Though there are no general-case weak-perspective invariants, there are special-case invariants of practical importance. The special-case weak-perspective invariants cited in the literature are derived from linear dependence relations and the invariance of this type of relation to linear transformation. The variation with respect to view is then studied for an important set of 2D line segment features: the relative orientation, size, and position of one line segment with respect to another. The analysis includes an important evaluation criterion for feature utility in terms of view-variation: the relationship between the fraction of views (over a view sphere) and the range of values assumed by a feature over these views. This relationship is a function of both the feature and the particular configuration of 3D line segments; an analysis and series of graphs are presented for each of the features and for different configurations of 3D line segments.

---

*This research was supported by the Defense Advanced Research Projects Agency under contract F30602-87-C-0140 and by Army ETL contract DACA76-89-C-0017.

1. Introduction

Familiar objects are typically three-dimensional structures sensed visually in the form of two-dimensional projections. While model-independent understanding of 3D structure [Marr82] is possible from motion, stereo, shading and texture, these cues may quite often be unavailable, unreliable or provide only a rough indication of the object structure. One of the outstanding problems in visual object recognition is the developement of systems capable of recognizing 3D objects by matching them to 2D image data.

The prediction-based methods developed in [Brooks81, Lowe85, Burns87, Korn87] constitute a promising approach to the recognition of 3D objects in 2D images. In this approach, recognition is achieved by (1) predicting characteristics of the object projections from all views, (2) matching these predictions against the input 2D image and (3) verifying all promising matches by determining the 3D pose of the object given the data matched. In its most general form a *prediction* expresses the expected values for a set of features of the object's projection. A *feature* can be any measurement or function of the projection, and the expectations can be any valid statement about the feature distribution.

Another fundamental problem in recognition is ensuring that computational costs for recognition grow only slowly with respect to model base size. This can be achieved by organizing the predictions across objects into a discrimination structure [Grimson84, Ikeuchi87]. Such a structure, called a *prediction hierarchy*, has been developed in our research on recognition systems [Burns87]. Objects with shared predictions are progressively discriminated by specifying additional features whose value distributions are different for the different objects. After the discrimination structure is compiled, the recognition system uses it to recursively match predictions that are progressively more object-specific.

For an image feature to be useful in discrimination, its distribution of values with respect to each object should be narrow, and the distributions with respect to different objects should be well separated. Since the camera viewpoint may be only partially constrained, the usefulness of a feature over the object projection is a function of how it varies with view. Hence, the property of *view-variation*, the extent to which a feature varies over given ranges of view, is fundamental to the recognition process. Since features of a projection usually "blow-up" to extreme values at some (usually small) set of views, a feature is considered here to have *low view-variation* if the variation is small in extent over a large fraction of the views [1]. Ideally, a feature would be *view-invariant*, that is, unaffected by change in view.

This paper presents a study of the variation of 2D point-set and line-segment features with respect to view. It is first established that no *general-case* view-invariant exists for any number of points, given true perspective, weak perspective or orthographic projection models. That is, there does not exist a feature that is view-invariant for all point sets of a given size n, for any n. Instead, there are only *special-case* invariants, features that are only view-invariant for special configurations of 3D points. It is important to determine the existence of general-case invariants since their distribution of values for each object would always have zero variance; only added noise and small separation in the value distribution for the different objects would hinder their usefulness. This is made clear in restricted recognition domains where special-case invariants become general-case for the domain. Some examples of this are the planar objects with 2D rigid transformations in [Tucker88], and planar objects with 3D rigid transformations and weak perspective projection in [Lamdan88a]. The property of measurement view-invariance can also be useful in other areas of vision, such as

[1] The term *quasi-invariant* [Binford87] has been used to denote something similar. Unfortunately, it is used to denote other things as well, and is thus avoided here.

the use of the cross-ratio in scene reconstruction [Mohr90]. For the domain of 2D projections of 3D points, the existence of general-case invariants does not seem to have been previously determined [Ahuja68, Duda73, Lowe85, Binford87, Aliomonos87, Lamdan88b, Weiss88]; in fact, in a chapter surveying projective geometry, Duda et al., explicitly mention the current lack of understanding on the subject.

The remainder of the paper then focuses on feature variation under weak perspective, a commonly used projection model in 3D recognition. Special-case invariants are surveyed for this projection model, and the variation with respect to view is studied for an important set of 2D line segment features: the relative orientation, size and position of one line segment with respect to another. The analysis includes an important evaluation criterion for feature utility in terms of view-variation: the relationship between fraction of views and the range of values assumed by a feature over these views. This relationship is a function of both the feature and the particular configuration of 3D line segments; an analysis and series of graphs are presented for each of the features and for different configurations of 3D line segments.

2. General-case view invariants

In this section, it is established that there is no feature of the perspective projection of n points, for any n, that is both a general-case view-invariant and non-trivial in the senses defined below. This result is then extended to orthographic and weak perspective projection models.

2.1 Definitions

The perspective object-to-image transformation is

$$\pi_{R,\vec{T}}(\vec{P}) = (F/z)\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}(R\vec{P} + \vec{T})$$

for 3D point[2] $\vec{P}$ at depth z, 3D rotation R, 3D translation $\vec{T}$ and focal length F. Each $(R, \vec{T})$ represents a distinct *view*.

We are considering features that are functions of the projections of 3D point sets and are defined almost everywhere. More precisely, a view is *degenerate* for a given feature and 3D point set if the feature is undefined for this view and point set. A *3D point set* is degenerate for a feature if the set of degenerate views for this feature and point set is not measure zero. We are considering features that have at most a measure zero collection of degenerate point sets.

The projection of the 3D point set S will be compactly represented as $\pi_{R,\vec{T}}(S)$, and the value for a given feature f, view $(R, \vec{T})$ and 3D point set S is $f(\pi_{R,\vec{T}}(S))$. Since we will be dealing with multiple point sets and these sets are ordered, we will refer to the jth point of the ith set as $\vec{P}_{i,j}$.

Def: A feature f is a *general-case* view-invariant for the class C^n of all 3D point sets of size n if, for all S in C^n and non-degenerate pairs of views $(R_1, \vec{T}_1, R_2, \vec{T}_2)$,

$$f(\pi_{R_1,\vec{T}_1}(S)) = f(\pi_{R_2,\vec{T}_2}(S)).$$

There is always a trivial f that can satisfy the above property: any constant feature,

[2]Henceforth, lower case signifies positions in the image coordinate frame, and upper case, positions in the object coordinate frame.

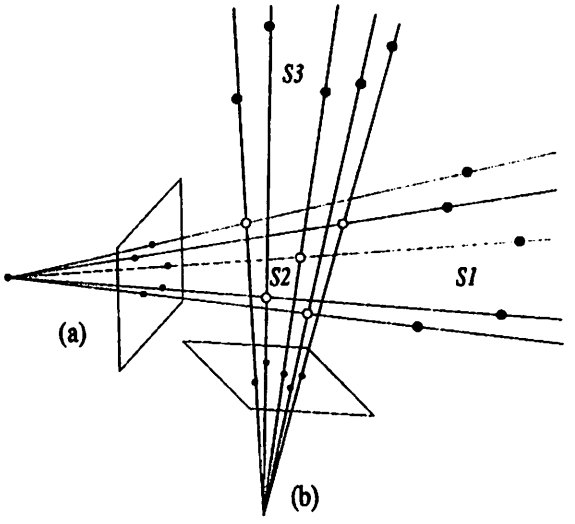


Figure 1: Perspective and projective correspondence. Point sets S_1 and S_2, and S_2 and S_3, are in perspective correspondence; point sets S_1 and S_3 are in projective correspondence through S_2.

where a *constant* feature has the same value for all 3D point sets and views. Thus, the following is another important property for a feature.

Def: A feature f is *non-trivial* if there exist two different, non-degenerate point-sets S_1 and S_2, and a pair of views $(R_1, \vec{T_1})$ and $(R_2, \vec{T_2})$, such that $f(\pi_{R_1,\vec{T_1}}(S_1)) \neq f(\pi_{R_2,\vec{T_2}}(S_2))$.

The following are two concepts from projective geometry [Duda73] that must be adapted to our problem domain of 3D point sets and 2D projections.

Def: *Perspective correspondence.* Two point sets of the same size are in perspective correspondence if there exist a pencil of rays and a one-one correspondence between the sets such that corresponding points are in the same ray (Figure 1a). In addition, all rays must be constrained to the same half-space to create an image by sectioning the pencil with a plane. Clearly, two such 3D point sets project to the same image if the camera focal point is the ray intersection point and the rays are sectioned by the image plane; in other words, $\pi_{R_1,\vec{T_1}}(S_1) = \pi_{R_2,\vec{T_2}}(S_2)$, for some $(R_1, \vec{T_1}, R_2, \vec{T_2})$.

Def: *Projective correspondence.* Two point sets are in projective correspondence if they can be connected by a chain of perspective correspondences. For example, in Figure 1, S_1 and S_3 are in projective correspondence through perspective correspondences with S_2.

2.2 Theorem and proof

Lemma 1 *A general-case view-invariant feature has the same value for point sets in projective correspondence if this correspondence is established through views that are not degenerate for the feature.*

Proof. For point sets in perspective correspondence S_1 and S_2, this is immediate. Being a general-case invariant, f has a single value f_1 for all projections of S_1, and similarly for S_2 (f_2). Since $\pi_{R_1,\vec{T}_1}(S_1) = \pi_{R_2,\vec{T}_2}(S_2)$, for some pair of non-degenerate views $(R_1, \vec{T}_1, R_2, \vec{T}_2)$, we have $f_1 = f(\pi_{R_1,\vec{T}_1}(S_1)) = f(\pi_{R_2,\vec{T}_2}(S_2)) = f_2$.

By definition, for point sets in projective correspondence S_1 and S_n, there exists a sequence of point sets $S_2, \ldots, S_{n-1}$ such that every pair (S_i, S_{i+1}) is in perspective correspondence, for $1 \leq i < n$. Given that these correspondences are established through non-degenerate views, by transitivity of equality, the feature values associated with S_1 and S_n must be equal. □

Lemma 2 *If two 3D point sets of the same size are identical for all but one point, then they are in projective correspondence through views that are not degenerate for a given feature, assuming the two point sets are not degenerate for this feature.*

Proof. Consider two such sets S_1 and S_2, and say that they differ in the position of the jth point; i.e., $\vec{P}_{1,j} \neq \vec{P}_{2,j}$ (Figure 2,a). In order to show *perspective* correspondence between

the sets, it is sufficient to show that there exists a pencil of rays that contains both sets. This can be done by placing the focal point (point of intersection for the rays) in the line that passes through $\vec{P}_{1,j}$ and $\vec{P}_{2,j}$. Clearly these two points will then have a ray in common and the other point pairs will share rays by virtue of the fact that they occupy the same positions in space.

Unfortunately, this perspective correspondence is established using a focal point selected from a restricted set, the line through $\vec{P}_{1,j}$ and $\vec{P}_{2,j}$, and it is possible that the views associated with this set are all degenerate for a given feature. For example, in Figure 2(b) this line contains another point $\vec{P}_{1,i}$ $(= \vec{P}_{2,i})$, and thus the projection of these points will be coincident for all views associated with the line, a degenerate condition for many features (Section 4 and 5).

However, a *projective* correspondence can always be established through non-degenerate views for a given feature. This can be done by considering a third, intermediate set of points S_3 (Figure 2,c) which is identical to both S_1 and S_2 except for point $\vec{P}_{3,j}$, which is positioned so that S_3 is in perspective correspondence with both S_1 and S_2 via non-degenerate focal points $\vec{F}_1$ and $\vec{F}_2$. From the construction (Figure 2,c), a non-degenerate projective correspondence can be established using a pair of points $(\vec{F}_1, \vec{F}_2)$ that satisfy the following three properties: (1) $\vec{F}_1$, $\vec{F}_2$, $\vec{P}_{1,j}$ and $\vec{P}_{2,j}$ are coplanar, (2) the lines through $(\vec{F}_1, \vec{P}_{1,j})$ and $(\vec{F}_2, \vec{P}_{2,j})$ are not parallel, and (3) $\vec{F}_1$ and $\vec{F}_2$ are not degenerate views for S_1 and S_2 respectively [3]. The first two properties are sufficient conditions for the existence of the intermediate point $\vec{P}_{3,j}$.

[3]These views must then also be non-degenerate for S_3, since this set is in perspective correspondence with S_1 and S_2 from these views.

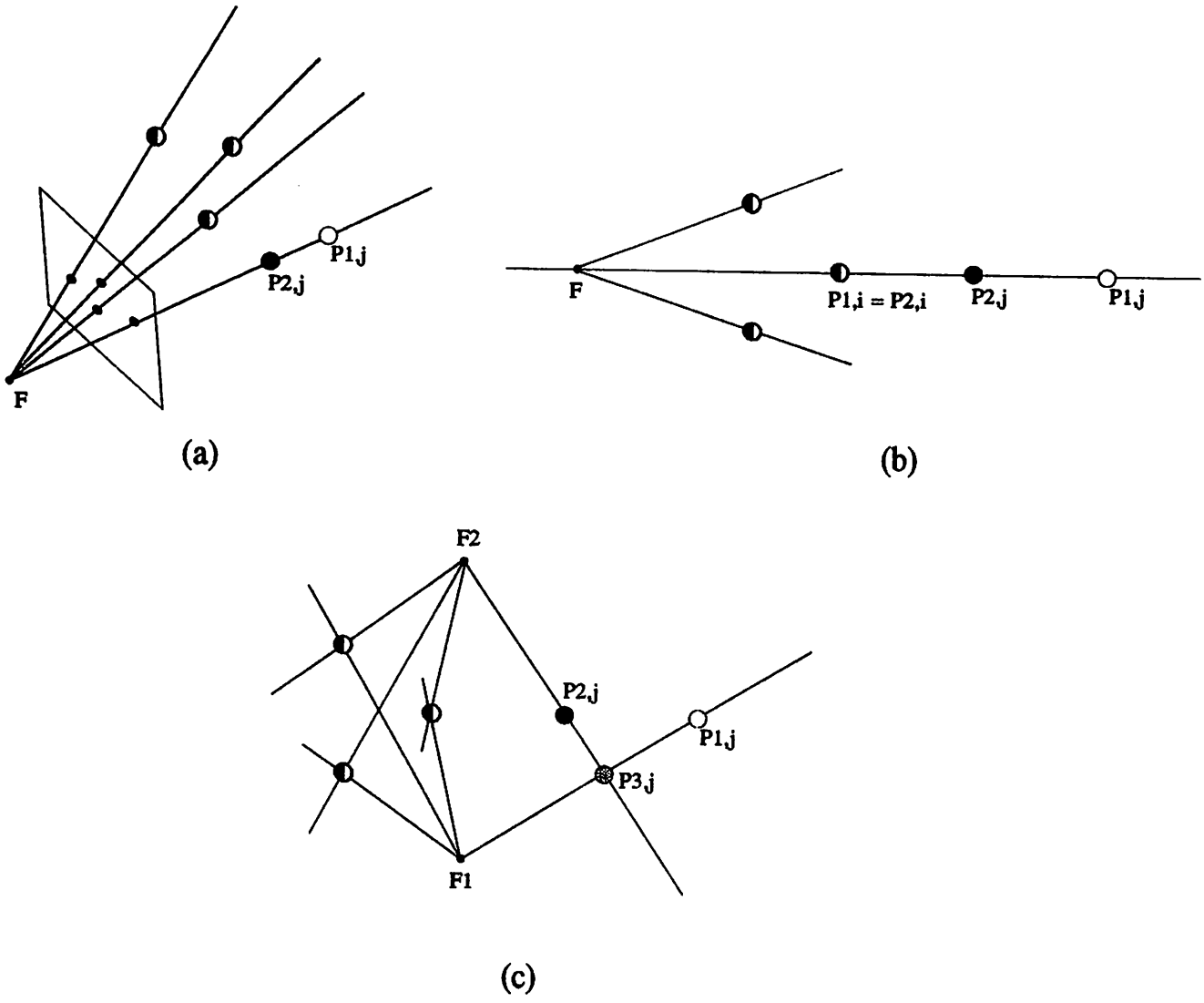


Figure 2: 3D point sets S_1 (white) and S_2 (black) are identical up to $\vec{P}_{1,j}$ and $\vec{P}_{2,j}$. (a) To show that such sets are in perspective correspondence select a focal point in the line through $\vec{P}_{1,j}$ and $\vec{P}_{2,j}$; in this way each pair of corresponding points, one from each set, share a ray. This correspondence cannot always be established using non-degenerate views, for example in (b), the projections of $\vec{P}_{1,j}$ and $\vec{P}_{1,i}$ will be coincident. (c) *Projective* correspondence through views non-degenerate for a given feature can be established using a set S_3, which is indentical to S_1 and S_2 for all but $\vec{P}_{3,j}$.

Given that there is only a measure zero set of degenerate views for any (non-degenerate) 3D point set and feature, it is always possible to find a pair of points that satisfy these conditions. Consider the pencil of planes that contain $\vec{P}_{2,j}$ and $\vec{P}_{1,j}$. At most a measure zero set of these planes can contain more than a measure zero set of degenerate $\vec{F}_1$, otherwise the total set of degenerate $\vec{F}_1$ in R^3 would not be of measure zero. Since the same situation holds for $\vec{F}_2$, and since the intersection of two measure zero sets is itself measure zero, at least one plane in the pencil contains at most a measure zero set of degenerate $\vec{F}_1$ and at most a measure zero set of degenerate $\vec{F}_2$. Consider this plane and a non-degenerate point $\vec{F}_1$ in it. Since there are non-degenerate $\vec{F}_2$ almost everywhere in the plane, there must be at least one not in the line through $\vec{P}_{2,j}$ that is parallel to the line $(\vec{F}_1, \vec{P}_{1,j})$. Hence, there is always a pair of points $(\vec{F}_1, \vec{F}_2)$ that satisfy the above three conditions, establishing the non-degenerate projective correspondence. □

Lemma 3 *All 3D point sets of size n are in projective correspondence through views that are not degenerate for a given feature, assuming the two point sets are not degenerate for this feature.*

Proof. Consider two sets of size n, S_1 and S_n. For these two sets, construct a sequence of sets $(S_2, \ldots, S_{n-1})$ such that adjacent pairs S_j and S_{j+1} are identical, except for their jth points, for $1 \leq j < n$. This sequence can be constructed by defining each S_j to be the union of the subset $\{\vec{P}_{1,i} | 1 \leq i < j\}$ of S_1 and the subset $\{\vec{P}_{n,i} | j \leq i \leq n\}$ of S_n.

From Lemma2, every adjacent pair of point sets in this sequence is in projective correspondence through views non-degenerate for any given feature, provided that these point sets are not themselves degenerate. By transitivity, the original point sets S_1 and S_n are

also in projective correspondence.

It is possible for some of the intermediate sets S_j to be degenerate for the given feature, even though S_1 and S_n are not. In these cases, we can select a set S_1' in projective correspondence with S_1 such that the sequence between S_1' and S_n contain no point sets degenerate for the feature. This is always possible since there can only be a measure-zero set of degenerate cases for the feature and any given set S_1 is always in projective correspondence, through non-degenerate views, with a non-measure-zero set [4] of S_1'. □

Theorem 1: There is no feature of n projected points, for any n, that is both a general-case view-invariant and non-trivial in the senses defined above.

Proof. Consider a general-case view-invariant f. For f to be non-trivial, there must exist at least one pair of point sets S_1 and S_2, such that $f(\pi_{R_1,\vec{T}_1}(S_1)) \neq f(\pi_{R_2,\vec{T}_2}(S_2))$ for some $(R_1, \vec{T}_1, R_2, \vec{T}_2)$.

Since all non-degenerate 3D point sets of the same size are in projective correspondence through non-degenerate views (Lemma 3) and every pair in projective correspondence through non-degenerate views must have the same value for any general-case view-invariant (Lemma 1), the necessary condition for a non-trivial feature cannot be established for any such general-case view-invariant. □

2.3 Extension to other projection models

Theorem 1 also holds for two other commonly used projection models: orthographic and weak perspective. Orthographic projection is the same as perspective, except that the

[4] This can be pictured by noting that, for a given ray bundle, each point in S_1 can move independently along its ray to produce a family of new sets S_1'. By repeatedly moving the points in this way, using non-degenerate and non-collinear ray intersection points, such a non-measure set of S_1' can be reached.

projection rays must always intersect at a point at infinity. Only Lemma 2 and Lemma 3 consider the projection geometry. In Lemma 2, since the construction step allows the intersection (focal) points to lie anywhere in given lines, they can always be placed at infinity. In addition, the existence of appropriate non-degenerate views can be demonstrated along the same lines as in the proof of Lemma 2. Lemma 3 is also valid for the orthographic case since any given point set S_1 is always in perspective correspondence with a non-measure-zero set of other point sets, even if the focal point is restricted to points at infinity.

Weak perspective is identical to orthographic except that the projection is also scaled by the average depth of the 3D point set. Since the added scaling is simply another degree of freedom in the transformation, the lack of general-case invariants under orthography must imply the same for weak perspective.

3. Weak perspective projection

The *weak perspective* projection model is used in our research for the purposes of predicting image properties. It is an approximation to perspective projection that is applied extensively in object recognition research as it simplifies the analysis and computation for 3D object recognition with reasonable results when the camera is far enough from the object relative to the depth variation of the object [Brooks81, Thompson87, Huttenlocher87, Lamdan88a]. Weak perspective is the same as the perspective projection defined in the last section, except that all of the points on a 3D object are treated as being at the same distance from the camera:

$$\pi_{R,\vec{T}}(\vec{P}) = (F/z_0)\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}(R\vec{P} + \vec{T})$$

for average distance z_0. Given this, the object-to-image transformation becomes much simpler: there is a single scale factor F/z_0 instead of a different variable z_i for each point, each with a non-linear effect. Another reason for working with a weak perspective model is that there are some useful special-case view-invariants (Section 4) that are approximately invariant for true perspective at an appropriately large camera distance. For example, parallelism can be effectively used as a line grouping criterion in grouping-based recognition [Lowe85]

It is important to note that prediction and match errors generated by assuming weak perspective have yet to be suitably analyzed. Some understanding can be gained by examining errors in the image point position. Consider a point in space $\vec{P}$ on an object with average camera distance z_0. For perspectively projected $\vec{p}$ and weak perspective approximation $\hat{p}$, we have

$$|\vec{p} - \hat{p}|/|\vec{p}| = \frac{\left| F(1/z - 1/z_0) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} (R\vec{P} + \vec{T}) \right|}{\left| F/z \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} (R\vec{P} + \vec{T}) \right|} = |z_0 - z|/z_0$$

Thus the proportional error in image position is equal to the relative depth $|z_0 - z|$ over the absolute average depth z_0. In [Thompson87], this depth ratio is recommended to be under one tenth for best results; at this range, the error in image position is at ten percent. This may be quite acceptable for recognition systems that use weak perspective to predict rough ranges in the orientation, size and position of image features. This is especially true if multiple features are used in discrimination and the objects are reasonably different. However, it is important to remember that a weak perspective approximation can contribute non-trivial error.

4. Special-case view-invariants under weak perspective

The weak perspective object-to-image transformation is a singular affine transformation that can be represented by a 2×3 rank 2 matrix and 2D translation vector [Lamdan88b]. By considering only point position *differences*, the translation component can be subtracted out and the vectors representing the 3D point position differences can be related to their projections through a linear transformation. This is important since the property of linear dependence for a set of vectors, the coefficients expressing the linear dependence, and the subspace dimension associated with the dependence are all invariant to linear transformation[5].

All of the weak perspective invariants discussed in the recognition literature [Brooks81, Lowe85, Lamdan88b] follow from this observation, and other, related special-case invariants can be deduced that may also be useful for recognition algorithms. For each type of special-case invariant, the required 3D point-set conditions can be specified as the linear dependence of some set of their position differences, and the invariant feature itself can be defined as a coefficient or subspace dimension associated with this linear dependence.

The different special-case conditions can be distinguished by two properties: the dimension of the subspace containing the set of 3D difference vectors, and the pattern of the point differences. The subspace dimensions given 3D point sets are two, one, and zero; the point difference patterns specify which pairs of points are being subtracted. Two types of difference patterns can be found in the invariants surveyed: differences from a single reference point $\vec{p}_1$ selected from the set, or $\{(\vec{p}_i - \vec{p}_1) | i \geq 2\}$, and differences between points that have been paired off, $\{(\vec{p}_{2i} - \vec{p}_{2i-1}) | i \geq 1\}$.

Table 1 shows the classification of special-case weak-perspective invariants based on the

[5]This follows from the fact that the null vector $\vec{0}$ is a fixed point for all linear transformations.

Subspace dimension	Pattern of 3D point differences	
	Differences from single reference point $\{(\vec{P}_i - \vec{P}_1) \mid i \geq 2\}$	Paired-off point differences $\{(\vec{P}_{2i} - \vec{P}_{2i-1}) \mid i \geq 1\}$
two	**Special-case constraint:** Four 3D points such that $\sum_{i=2}^{4} c_i(\vec{P}_i - \vec{P}_1) = \vec{0}$ for non-zero c_i. (Four coplanar points.) **View invariants:** (1) $(c_3/c_2, c_4/c_2)$. Example of use: the affine coordinate features in [Lamdan88a] (2) Subspace dimension of dependence is 2, which is not significant for 2D images.	**Special-case constraint:** Six 3D points such that $\sum_{i=1}^{3} c_i(\vec{P}_{2i} - \vec{P}_{2i-1}) = \vec{0}$ for non-zero c_i. (The points need not be coplanar.) **View invariants:** (3) $(c_2/c_1, c_3/c_1)$. No known example of use in recognition literature. (4) Subspace dimension of dependence is 2, which not significant for 2D images.
one	**Special-case constraint:** Three 3D points such that $\sum_{i=2}^{3} c_i(\vec{P}_i - \vec{P}_1) = \vec{0}$ for non-zero c_i. (Three collinear points.) **View invariants:** (5) c_3/c_2. Ratio of distances along line, example of use: the *approach ratio* in [Brooks81]. (6) Subspace dimension of dependence is 1, collinearity.	**Special-case constraint:** Four 3D points such that $\sum_{i=1}^{2} c_i(\vec{P}_{2i} - \vec{P}_{2i-1}) = \vec{0}$ for non-zero c_i. (Endpoints of two parallel line segments.) **View invariants:** (7) c_2/c_1. Ratio of distances measured in parallel directions. (8) Subspace dimension of dependence is 1, parallelism. Example of use: grouping-based recognition in [Lowe85].
zero	**Special-case constraint:** Two 3D points such that $c_2(\vec{P}_2 - \vec{P}_1) = \vec{0}$ for non-zero c_2. (Two coincident points.) **View invariants:** (9) c_2. Uninteresting (can be anything but zero). (10) Subspace dimension of dependence is 0, coincidence. Example of use: coincidence of line segment endpoints for grouping-based recognition in [Lowe85].	Same case as left panel.

Table 1: Classification of special-case view-invariants under weak perspective. All invariants discussed in the literature are functions of point position differences, and they can be distinguished by two properties: the dimension of the subspace containing the set of 3D difference vectors (row), and the pattern of the point differences (column).

above two properties. It is interesting to note that special-case invariants that have not been found in the recognition literature can be deduced from this framework; see for example table entry (3), the invariants classified under (*dimension two, paired-off point differences*). Many of the special-case invariants represented in the table can be equivalently expressed as special cases for one of the four features analysed in the next section; this will be made clear in that section.

There are also special-case invariants under perspective projection [Duda73, Lamdan88b] analogous to those under the column *Differences from single reference point* in Table 1. The perspective invariant called the *cross-ratio* is analogous to the ratio of coefficients (table entry 5), but requires four points instead of three, and *projective coordinates* correspond to the affine coordinates (table entry 1), but require five points instead of four. There are also projective invariants of planar curves that can be used in recognition [Forsyth90]. There do not seem to be special-case invariants under perspective projection analogous to the ones in the second column (table entries 3, 7, 8), utilizing linear dependence of paired-off point differences.

5. View variation of relative orientation, size and position

Though view-invariance is restricted to certain special 3D line segment configurations, the property of low view-variation, as defined in Section 1, is much more common. It is also of practical importance, since a reasonably small variation in feature value for each object, relative to the value variation *across* the objects, can be quite effective for 3D object discrimination in the sense discussed in Section 1. It is difficult to be precise about what we mean by *low* view-variation as it is context dependent, being meaningful in terms of the

feature value distributions over the particular set of objects being discriminated. However, we can be precise about the extent of variation of a feature as a function of particular 3D line segment configurations and portions of view space.

This section presents an analysis of view-variation for four important features of a projected line segment pair. They are the *relative* orientation, size and position of one segment with respect to the other. These features are defined and motivated in the next subsection. For each feature, the following results are presented:

- *The feature value as a function of view and 3D line segment configuration.* The value is expressed as a function only of the view parameters and parameters of the 3D segment configuration that affect it. The expression is simplified as much as possible by using an appropriate object coordinate frame and coordinate transformation representation.

- *Qualitative analysis.* This includes a description of the 3D line segments and ranges of views for which the feature variation is most and least constrained, and what the feature values are at these points.

- *Quantitative analysis.* Graphs are presented and discussed showing the feature variation as a function of view and 3D segment configuration. In addition, the relationship between extent of view-variation, the 3D segments and the range of views is explicitly presented. Feature interval size is plotted against view region size for various 3D segments.

5.1 The features

The view-variation of four features are studied in this section: the orientation, size and position (two components) of one projected line segment with respect to another (Figure 3a).

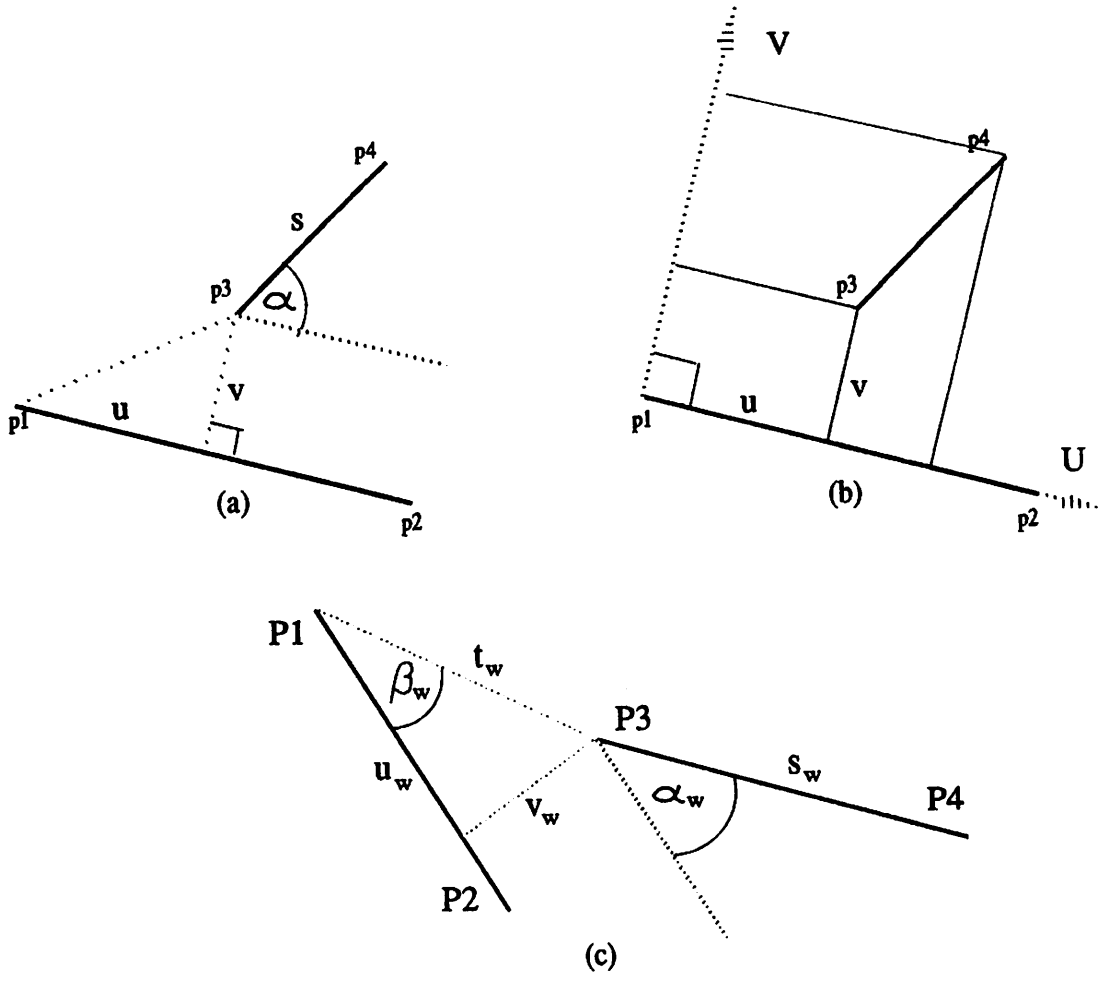


Figure 3: The relative features of line segments: (a) image segments (bold) and the projected relative orientation α, size s and position (u, v), (b) coordinate frame (U, V) aligned with first segment and used to measure second segment, (c) corresponding 3D segments and features.

These *relative* features can be pictured in the following way (Figure 3b). First, a coordinate frame is defined in terms of the first projected segment: the origin is set at its first endpoint, one axis (U) is aligned with the segment, the other is orthogonal to it (V), and the scale is such that the first segment is of unit length. The orientation, position and size of the second segment is then measured with respect to this coordinate frame. Specifically, given the projected segments ($\vec{p_1}$, $\vec{p_2}$) and ($\vec{p_3}$, $\vec{p_4}$), the four features are the defined in the following manner:

- relative orientation α. This is measured counterclockwise from $(\vec{p_2} - \vec{p_1})$ to $(\vec{p_4} - \vec{p_3})$ and has a magnitude of $\arccos[(\vec{p_2} - \vec{p_1}) \cdot (\vec{p_4} - \vec{p_3})/|\vec{p_2} - \vec{p_1}||\vec{p_4} - \vec{p_3}|]$,
- relative size s is the length ratio, $|\vec{p_4} - \vec{p_3}|/|\vec{p_2} - \vec{p_1}|$,
- relative position (u, v) is the position of endpoint $\vec{p_3}$ relative to the segment $(\vec{p_1}, \vec{p_2})$. It is the displacement of $\vec{p_3}$ from $\vec{p_1}$ measured along and normal to $(\vec{p_2} - \vec{p_1})$ and divided by its length, or $u = (\vec{p_3} - \vec{p_1}) \cdot (\vec{p_2} - \vec{p_1})/|\vec{p_2} - \vec{p_1}|^2$ and $v = (\vec{p_3} - \vec{p_1}) \cdot (\vec{p_2} - \vec{p_1})^{\perp}/|\vec{p_2} - \vec{p_1}|^2$, where $(\vec{p_2} - \vec{p_1})^{\perp}$ is $(\vec{p_2} - \vec{p_1})$ rotated by 90 degrees in the image plane.

For each of the relative 2D features of the projected segments there is a corresponding parameter of the 3D segment configuration being projected; these 3D world parameters are indicated by: α_w, s_w, u_w, v_w (Figure 3c). In addition, to analyze (u, v), it is useful to keep in mind β_w, the angle between $(\vec{P_2} - \vec{P_1})$ and $(\vec{P_3} - \vec{P_1})$, and t_w, the distance between $\vec{P_3}$ and $\vec{P_1}$ scaled by the length of $(\vec{P_2} - \vec{P_1})$.

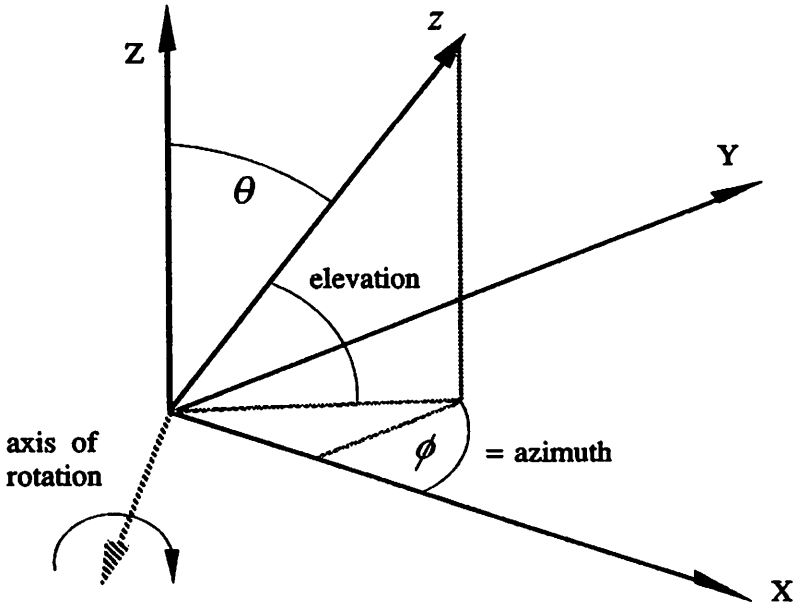


Figure 4: View sphere coordinates (ϕ, θ) expressing the orientation of optical z axis relative to object Z (north pole). The angle ϕ is the azimuth (angle about the pole) and θ is the angle from the pole (90 - elevation).

5.2 View representation

Each feature value is a function of view and the 3D segments; however only certain parameters of the view and 3D segment configuration affect this value. In this section, we define and justify a simplified expression of the weak perspective object-to-image transformation that is only in terms of the view parameters relevant to the discussed features.

Given weak perspective projection, all of the features considered here are invariant to rotation about the optical z axis and all three degrees of freedom in translation. Therefore, the only aspect of the view transformation that affects these features is the orientation of the optical z axis relative to the object coordinate system, say the object Z axis. The relative orientation of z can be expressed in terms of two parameters (ϕ, θ), which represent the positions on a unit *view* sphere about the object with the north pole at Z (See [Horn86] and

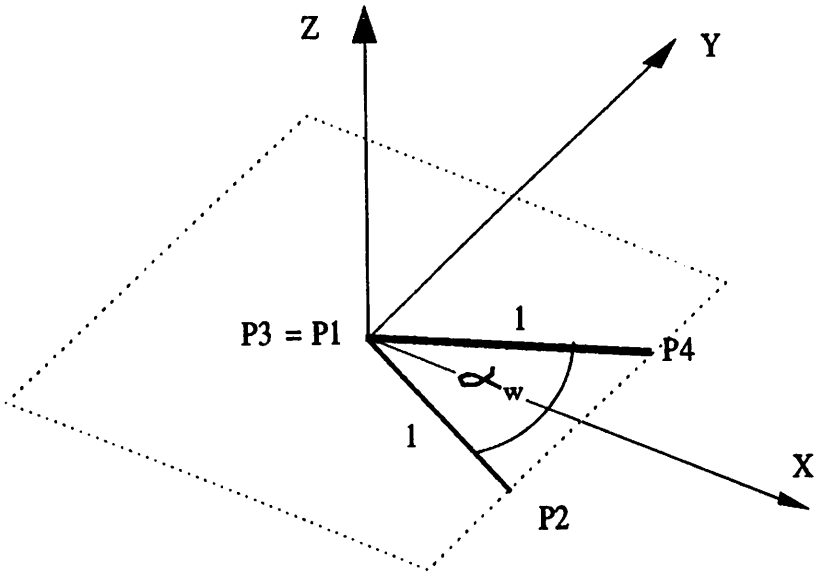


Figure 5: 3D line segment configuration with interior angle α_w. The view variation of α and s can be studied by analyzing the projections of this pair.

Figure 4). The angle ϕ is the azimuth (angle about the pole) and θ is the angle from the pole (90 - elevation). The optical z axis in object coordinates is then $(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Using this representation, the essential rotation from object to image can be expressed in terms of the two view sphere coordinates. The magnitude of the rotation from Z to z is simply θ, and the axis of the rotation $z \times Z/|z \times Z| = (\sin\phi, -\cos\phi, 0)$. We will refer to this rotation as $R_{\phi,\theta}$. The object-to-image transformation under consideration can now be expressed as $\vec{p} = \pi R_{\phi,\theta}\vec{P}$, where π is orthographic projection.

5.3 Relative orientation, α

In this section we first express the relative orientation of two projected segments as a function of (ϕ, θ) and the relevant parameters of the 3D segments. The magnitude of α is simply the angle between the two projected segments or $\arccos|(\vec{p_2} - \vec{p_1}) \cdot (\vec{p_4} - \vec{p_3})/|\vec{p_2} - \vec{p_1}||\vec{p_4} - \vec{p_3}||$ for projected endpoints $\vec{p_i}$.

The feature α is strictly a function of the projected differences $(\vec{p_4} - \vec{p_3})$ and $(\vec{p_2} - \vec{p_1})$, and, for weak perspective projection, these differences are unaffected by translation of the 3D segments relative to the object center and to each other. Also, the angle between the projected segments is unaffected by the lengths of the 3D segments. Thus, the only two factors affecting α are the orientation of the segments with respect to the camera (ϕ, θ) and the angle between them α_w. Thus, without loss of generality, we can analyze the variation of α by studying the projections of the 3D segment pair (see Figure 5):

$$\begin{aligned}
\vec{P_1} &= \vec{P_3} = (0,0,0), \\
\vec{P_2} &= (\cos(\alpha_w/2), -\sin(\alpha_w/2), 0), \\
\vec{P_4} &= (\cos(\alpha_w/2), \sin(\alpha_w/2), 0),
\end{aligned}$$

where α_w is the angle between the 3D segments. The feature α can then be expressed as a function of the view parameters (ϕ, θ) defined in Section 5.2 and the 3D angle α_w:

$$\begin{aligned}
\alpha &= \arccos(\vec{p_2} \cdot \vec{p_4} / |\vec{p_2}||\vec{p_4}|), \\
\vec{p_2} &= \pi R_{\phi,\theta}(\cos(\alpha_w/2), -\sin(\alpha_w/2), 0)^T \\
\vec{p_4} &= \pi R_{\phi,\theta}(\cos(\alpha_w/2), \sin(\alpha_w/2), 0)^T
\end{aligned}$$

Figure 6(a,c,e,g) shows polar plots of α as a function of view (ϕ, θ), given the above 3D line segments with (a) $\alpha_w = 90$ degrees, (c) 45 degrees, (e) 22.5 degrees, and (g) 11.25 degrees. The plots are over a hemisphere of views, where ϕ ranges from 0 to 360 and θ, from 0 to 90. The angle ϕ is represented by the clockwise angle about the center of the plot, with $\phi = 0$ at the three o'clock position, and θ is represented by the radius out from the center. The contour lines are constant values of α (slices of the surface at various α). A

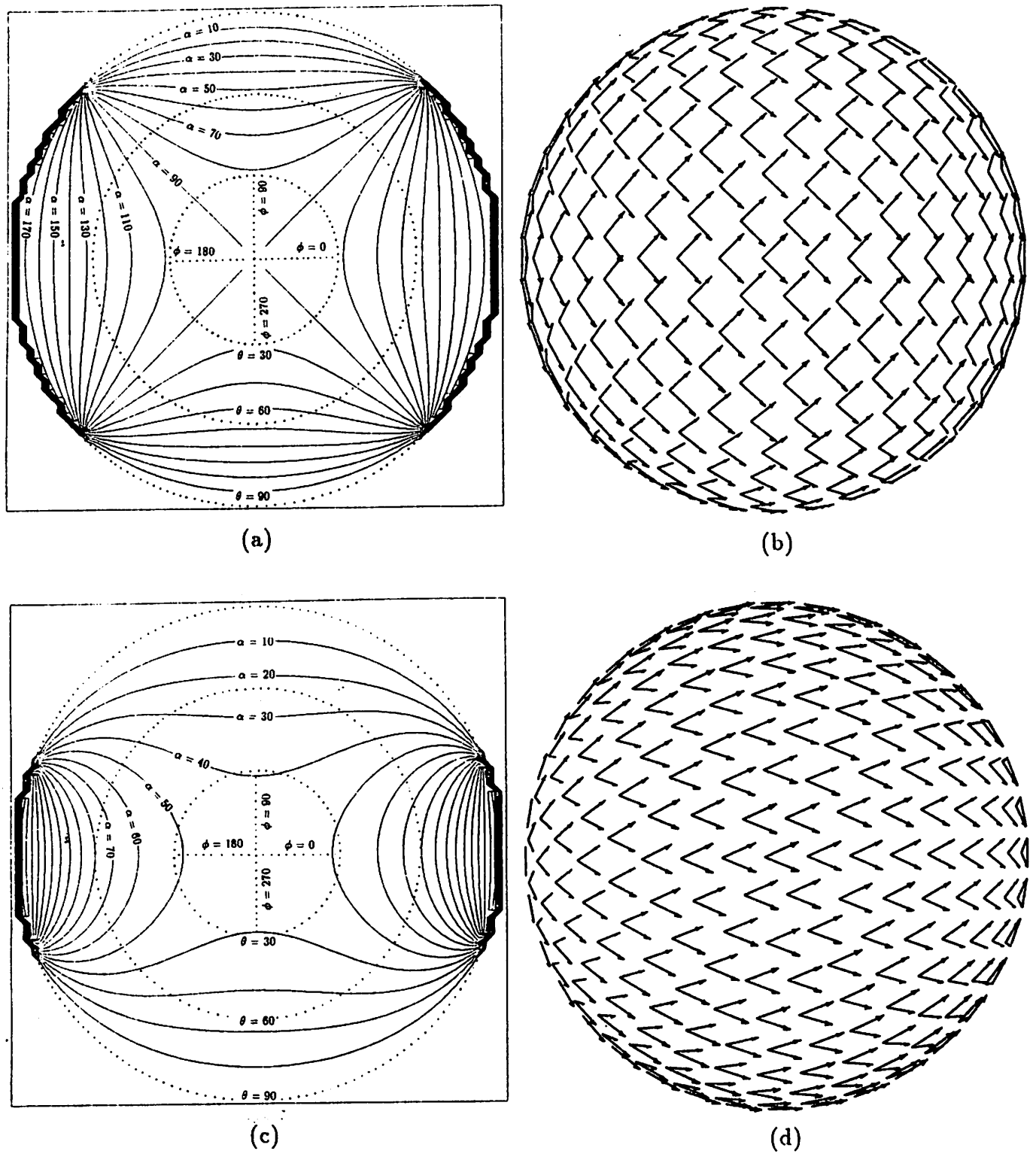


Figure 6: Polar plots of α as a function of view (θ, ϕ), for (a) 3D angle $\alpha_w = 90$ and (c) 45 degrees. The angle ϕ is represented by the clockwise angle about the center of the plot, with $\phi = 0$ at the three o'clock position, and θ is represented by the radius out from the center. The contour lines are constant values of α (slices of the surface at various α). Also shown are projections of the 3D line segments in Figure 5 from different views and for (b) 3D angle $\alpha_w = 90$ and (d) 45 degrees. Each projection is plotted on the sphere at roughly the view position (θ, ϕ) from which the segments would be seen this way. The sphere is orthographically projected, with the object Z is pointing out of the page and the X axis pointing to the left.

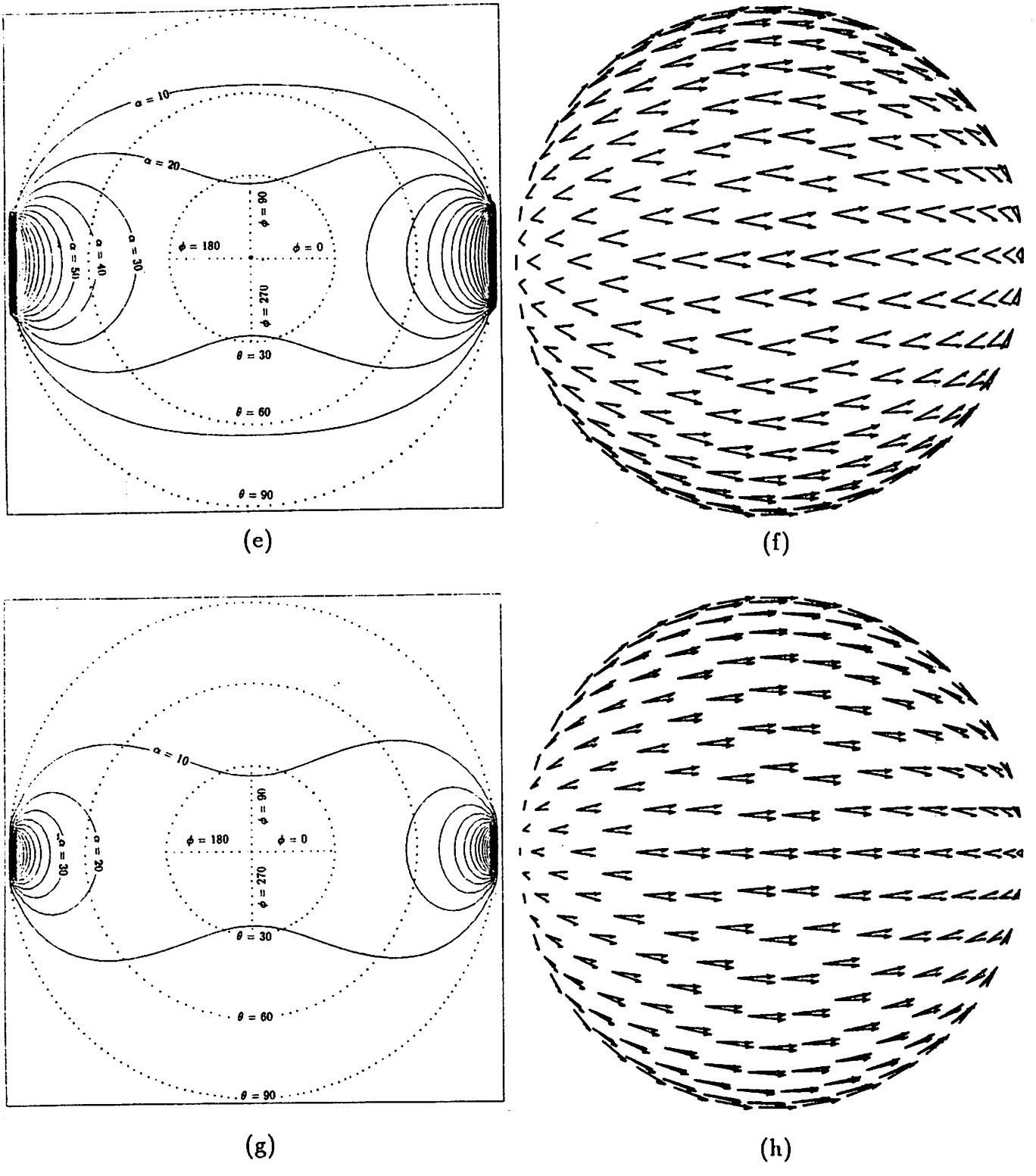


Figure 6: (Continued.) Polar plots of α as a function of view (θ, ϕ), for (e) 3D angle $\alpha_w = 22.5$ and (g) 11.25 degrees. Also shown are projections of the 3D line segments in Figure 5 from different views and for (f) 3D angle $\alpha_w = 22.5$ and (h) 11.25 degrees.

sense of where the variation is greatest, and by how much, can be gained by observing the contour line density: the denser the lines, the greater the variation in feature α with respect to change in view (ϕ, θ). For example, the variation is greatest when the view direction approaches the orientation of the 3D line segments (bold lines) and slowest when oriented normal to the plane containing the line segments (xy plane).

Figure 6(b,d,f,h) shows a less quantitative but more intuitive picture of the view variation for the same 3D angles. In this figure, the actual line segment projections are shown for various views about the sphere parametrized by (θ, ϕ). Each projection is plotted on the sphere at the view position from which it would be seen this way, and α is the angle between the projected lines at that view position. The sphere is orthographically projected, with the object Z pointing out of the page and the X axis pointing to the left.

From the equations and plots, the following observations can be made:

- When α_w is zero, then α is constant (zero) for all (θ, ϕ). This occurs exactly when the 3D segments are parallel, or equivalently, when the point differences $(\vec{P}_2 - \vec{P}_1)$ and $(\vec{P}_4 - \vec{P}_3)$ are linearly dependent. In this situation, α is equivalent to the special case invariant represented in Table 1, entry (8).

- As α_w approaches zero, the variation as a function of (θ, ϕ) is slow over most of the views.

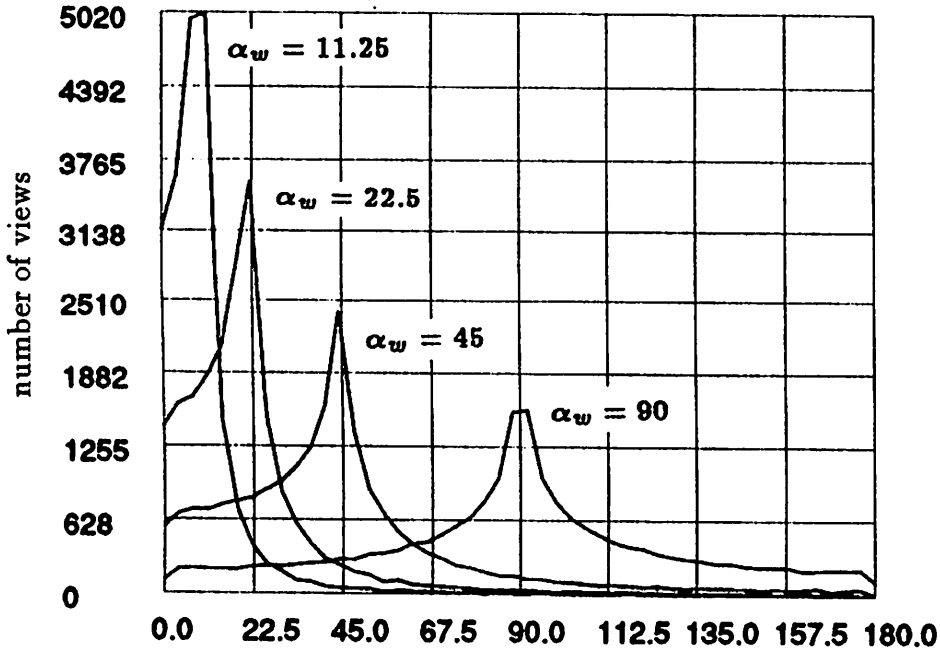
- When $\alpha_w > 0$, there always exists some view where $\alpha = 180$ degrees and $\alpha = 0$. In other words, α covers the full range of possible values. Angle α approaches 0 when (θ, ϕ) approach (90, 90) and (90, -90), and α approaches 180 when (θ, ϕ) approach (90, 0) and (90, 180).

- α is slowest in variation when the view is most parallel to the Z axis and normal to the plane containing the segments ($\theta = 0$). At this point $\alpha = \alpha_w$. The variation in α is still fairly slow for views near the object Z axis, especially if the 3D angle α_w is reasonably small.

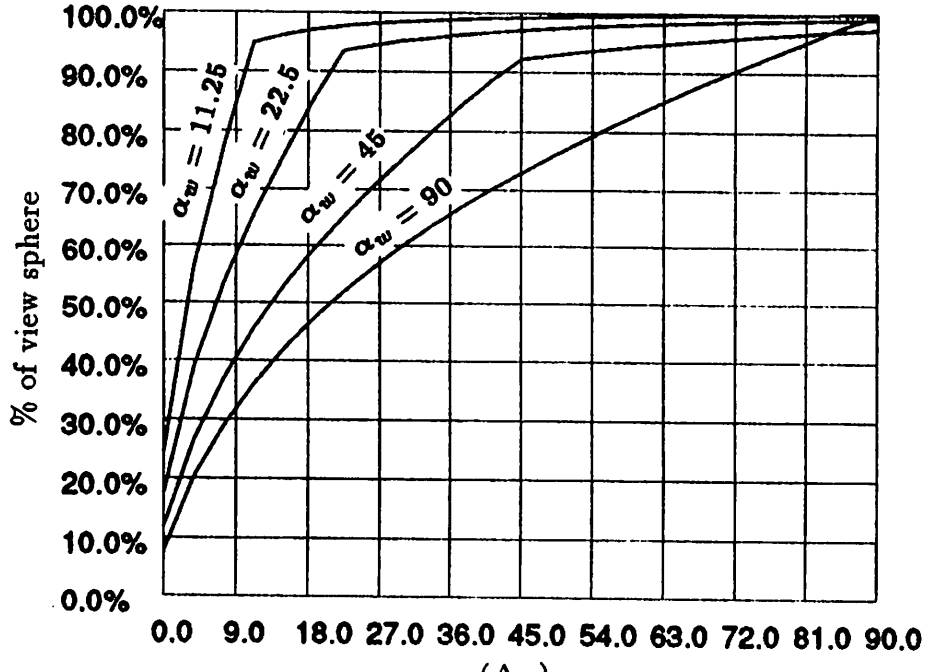
A more quantitative picture can be gained by studying the distribution of α as (ϕ, θ) varies, for the different 3D angles α_w. Using an approximately regular sampling of the sphere of viewpoints parametrized by (ϕ, θ), histograms of the number of views that fall within regular intervals of α were made for various 3D angles α_w (Figure 7a). For example, the wider, symmetric graph represents the distribution of α over all views for 3D line segments with angle $\alpha_w = 90$ degrees. The projected feature α ranges from 0 to 180 degrees, and for the histograms, this range was divided into fifty regular intervals. The view-sphere sampling was done so that the views are approximately one degree apart [6].

As the histograms show, the distributions of the projected features are strongly concentrated about the value of the true 3D angles, α_w, and this concentration becomes more pronounced as α_w approaches zero (similarly for 180 degrees). An even more revealing picture can be gained by counting the number of views for which α falls within some interval about the 3D angle α_w; i.e., $[\alpha_w - \Delta\alpha,\ \alpha_w + \Delta\alpha]$. Figure 7(b) shows the percentage of the view-sphere that falls within this interval for $\Delta\alpha$ ranging from 0 to 90 degrees and α_w of 90, 45, 22.5 and 11.25 degrees. The amount of the view sphere that falls within an interval clearly gets larger as α_w decreases. For the values of α_w studied, a smallish interval of $[\alpha_w - 15,\ \alpha_w + 15]$ covers approximately 42, 53, 75 and 96 percent of the sphere, respectively.

[6] The actual sampling produced 20447 samples for a hemisphere, with an average angle between adjacent ones of 0.9999 degrees and standard deviation of 0.001 degrees. (For discussion on the sampling technique, see Appendix)



(a)



(b)

Figure 7: Quantitative picture of α distribution. (a) Histograms of the number of views that fall within regular intervals of α for a hemisphere of views and for objects with various true 3D angles, α_w. (The range of α is divided into fifty intervals.) (b) The percentage of the view sphere that falls within the interval $[\alpha_w - \Delta\alpha, \alpha_w + \Delta\alpha]$ for various $\Delta\alpha$ ranging from 0 to 90 (degrees) and α_w of 90, 45, 22.5 and 11.25 (degrees).

To guarantee that most of the sphere (say 80 percent of it) falls within the interval, $\Delta\alpha$ has to be approximately 55, 34, 16 and 8 degrees for each of the α_w, respectively.

5.4 Relative size, s.

Relative size, s, is the ratio of the two projected segment lengths or, $s = |\vec{p_4} - \vec{p_3}|/|\vec{p_2} - \vec{p_1}|$. This feature is affected by the same view and 3D line segment structure parameters as α, for the same reasons, except that s is also clearly affected by the ratio of the lengths of the 3D segments, s_w. Thus, 2D feature s is strictly a function of view (ϕ, θ), the angle between the 3D line segments α_w and s_w. Without loss of generality, this function can be represented by considering the same 3D line segment configuration as for α, except that $\vec{P}_4$ is scaled by s_w. By considering the projections of this configuration, we get the following expression for s:

$$\begin{aligned} s &= |\vec{p_4}|/|\vec{p_2}|, \\ \vec{p_2} &= \pi R_{\phi,\theta}(\cos(\alpha_w/2), -\sin(\alpha_w/2), 0)^T \\ \vec{p_4} &= s_w \pi R_{\phi,\theta}(\cos(\alpha_w/2), \sin(\alpha_w/2), 0)^T \end{aligned}$$

From the above equations, s is simply a linear function of s_w. The variation of s as the view varies, for different α_w, can be appreciated by observing the ratio of lengths of the projected line segments in Figure 6(b,d,f,h). Contour graphs of s as a function of view (ϕ, θ), for different α_w, is shown in Figure 8. From the equations and graphs, it is clear that:

- When α_w is zero then s is constant (s_w) for all (θ, ϕ). This occurs exactly when the 3D segments are parallel, or equivalently, when the point differences $(\vec{P}_2 - \vec{P}_1)$ and $(\vec{P}_4 - \vec{P}_3)$ are linearly dependent. In this situation, s is equivalent to the special case invariant represented in Table 1, entry (7).

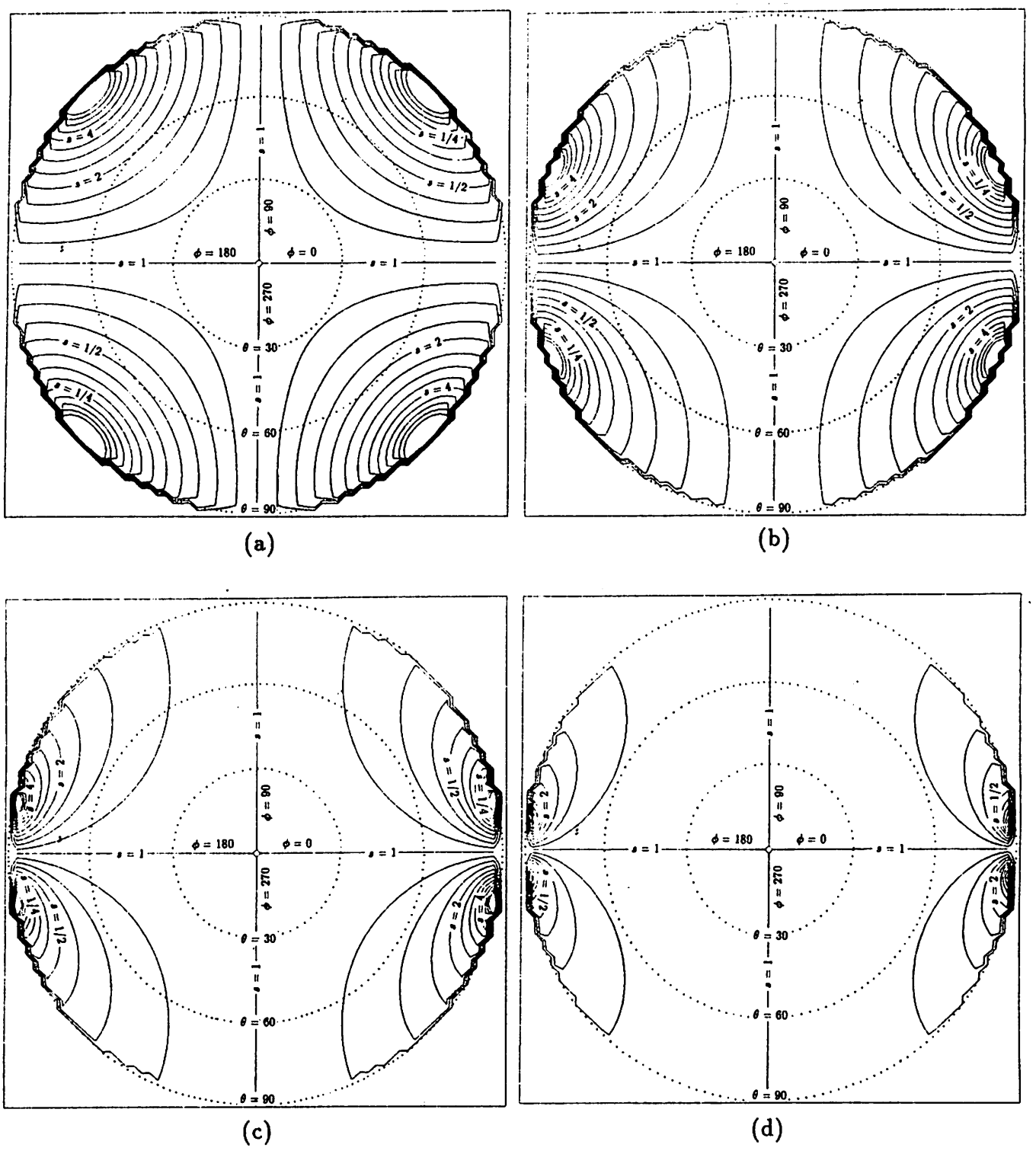


Figure 8: Plots of s as a function of θ and ϕ, for $s_w = 1$ and different values of α_w. The angle from each plot x axis represents ϕ (*azimuth*) and the radius from the center represents θ ($90 - elevation$). The contour lines indicate different values of s. The angle (in degrees) between the 3D segments, α_w, is (a) 90, (b) 45, (c) 22.5 and (d) 11.25.

- As α_w approaches zero, the variation as a function of (θ, ϕ) is slow over most of the views.

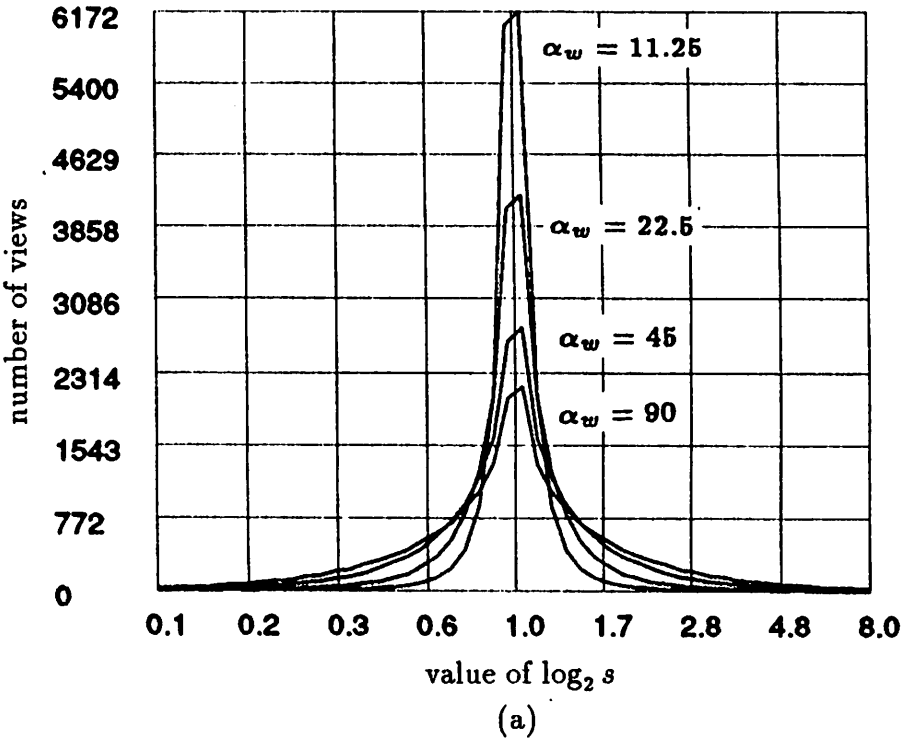
- When $\alpha_w > 0$, there always exists some view where $s = \infty$ and $s = 0$. In other words, as in the case of α, s covers the full range of possible values. The value s approaches 0 when (θ, ϕ) approach $(90, \alpha_w/2)$ and $(90, 180 + \alpha_w/2)$. This is when the view aligns with the second segment. It approaches ∞ when (θ, ϕ) approach $(90, -\alpha_w/2)$ and $(90, 180 - \alpha_w/2)$ - when the view lines up with the first segment.

- As in the case of α, s is slowest in variation when the view is most normal to the plane containing the segments ($\theta = 0$). At this point $s = s_w$. Also, variation in s is still fairly slow for views near the Z axis, especially if the 3D angle α_w is reasonably small.

Figure 9(a) shows the histograms of the number of views that fall within regular intervals of s for various 3D angles of α_w, with s_w set to one and the same view sampling scheme as in Section 5.3. The projected size ratio s actually ranges from zero to infinity - but for these examples, the histograms range from zero to eight and contain almost all of the 20,477 samples. For reasons of symmetry, the s axis is plotted using a $\log_2$ scale[7].

As the histograms show, the distributions of the projected features are strongly concentrated about the value of the true 3D length ratio (in this case $s_w = 1$), and this concentration becomes more pronounced as α_w approaches zero (similarly for 180 degrees). A more revealing picture can be gained by counting the number of views from which s falls within some interval about the true feature $s_w = 1$. Again, for reasons of symmetry, the interval is taken to be $[1/\Delta s, \Delta s]$.

[7]In this way, the length ratio $s = |\vec{p_1}|/|\vec{p_2}|$ has the same variational behavior as its inverse, which seems reasonable.



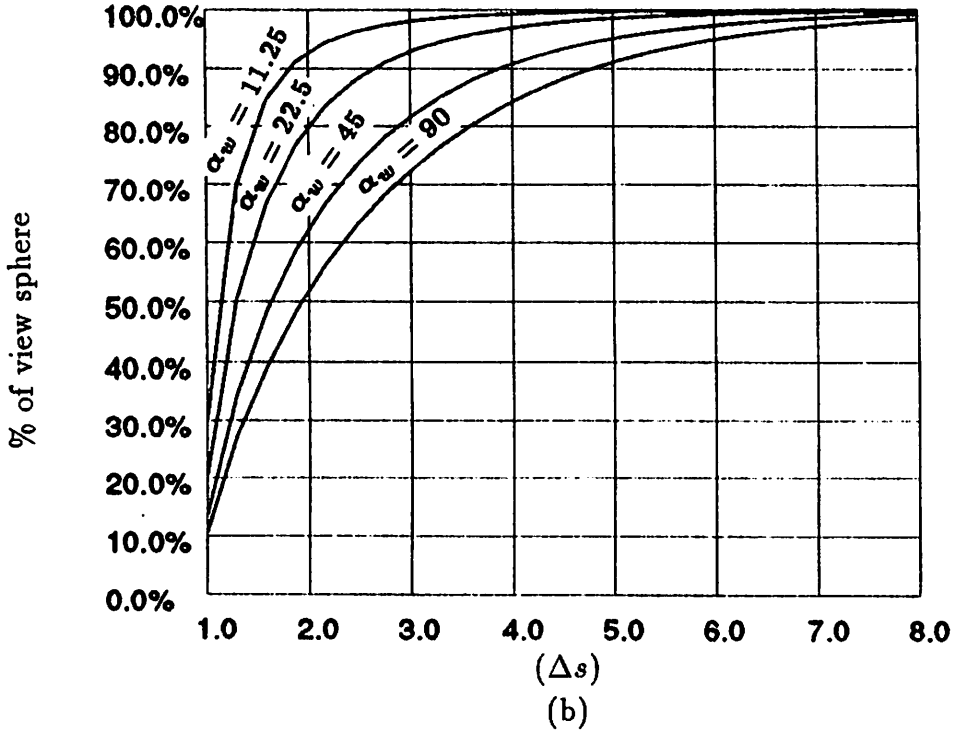


Figure 9: Quantitative picture of s distribution. (a) Histograms of the number of views that fall within regular intervals of $\log_2 s$ for a hemisphere of views and for objects with various true 3D angles, α_w (and true $s_w = 1$). (The range of s is divided into fifty intervals of $\log_2$ spacing and the plot uses a $\log_2$ scale for s.) (b) The percentage of the view sphere that falls within the interval $[1/\Delta s, \Delta s]$ for various Δs ranging from 1 to 8 and α_w of 90, 45, 22.5 and 11.25 (degrees).

Figure 9(b) shows the percentage of the view-sphere that falls within this interval for Δs ranging from 1 to 8 and α_w of 90, 45, 22.5 and 11.25 degrees. The amount of the view sphere that falls within an interval clearly gets larger as α_w decreases. For the values of α_w studied, the smallish interval of [1/2, 2] covers approximately 53, 64, 80 and 93 percent of the sphere respectively. To guarantee that most of the sphere (say 80 percent of it) falls within the interval, Δs has to be approximately 3.6, 2.8, 2 and 1.5 for decreasing values of α_w respectively.

5.5 Relative position, (u, v)

Relative position comes in two components (u, v) measured in orthogonal directions in the image plane (Section 5.1).

$$u = (\vec{p_3} - \vec{p_1}) \cdot (\vec{p_2} - \vec{p_1}) / |\vec{p_2} - \vec{p_1}|^2,$$

$$v = (\vec{p_3} - \vec{p_1}) \cdot (\vec{p_2} - \vec{p_1})^{\perp} / |\vec{p_2} - \vec{p_1}|^2,$$

where $(\vec{p_2} - \vec{p_1})^{\perp}$ is $(\vec{p_2} - \vec{p_1})$ rotated by 90 degrees in the image plane.

The 2D relative position features are functions of the vectors $(\vec{p_2} - \vec{p_1})$ and $(\vec{p_3} - \vec{p_1})$, instead of $(\vec{p_2} - \vec{p_1})$ and $(\vec{p_4} - \vec{p_3})$. This means that the situation is exactly the same as for the 2D feature s except that u and v are functions of the view (ϕ, θ), the 3D angle β_w (instead of α_w) and the 3D length ratio t_w (instead of s_w), for β_w and t_w defined in 5.1. Without loss of generality, these functions can be represented by considering a 3D line segment configuration equivalent to the one used for s, except that $\vec{P}_3$ is in the position of $\vec{P}_4$, and the relevant angle and length are β_w and t_w. By considering the projections of this configuration, we get the following expressions for u and v:

$$u = \vec{p_3} \cdot \vec{p_2} / |\vec{p_2}|^2,$$

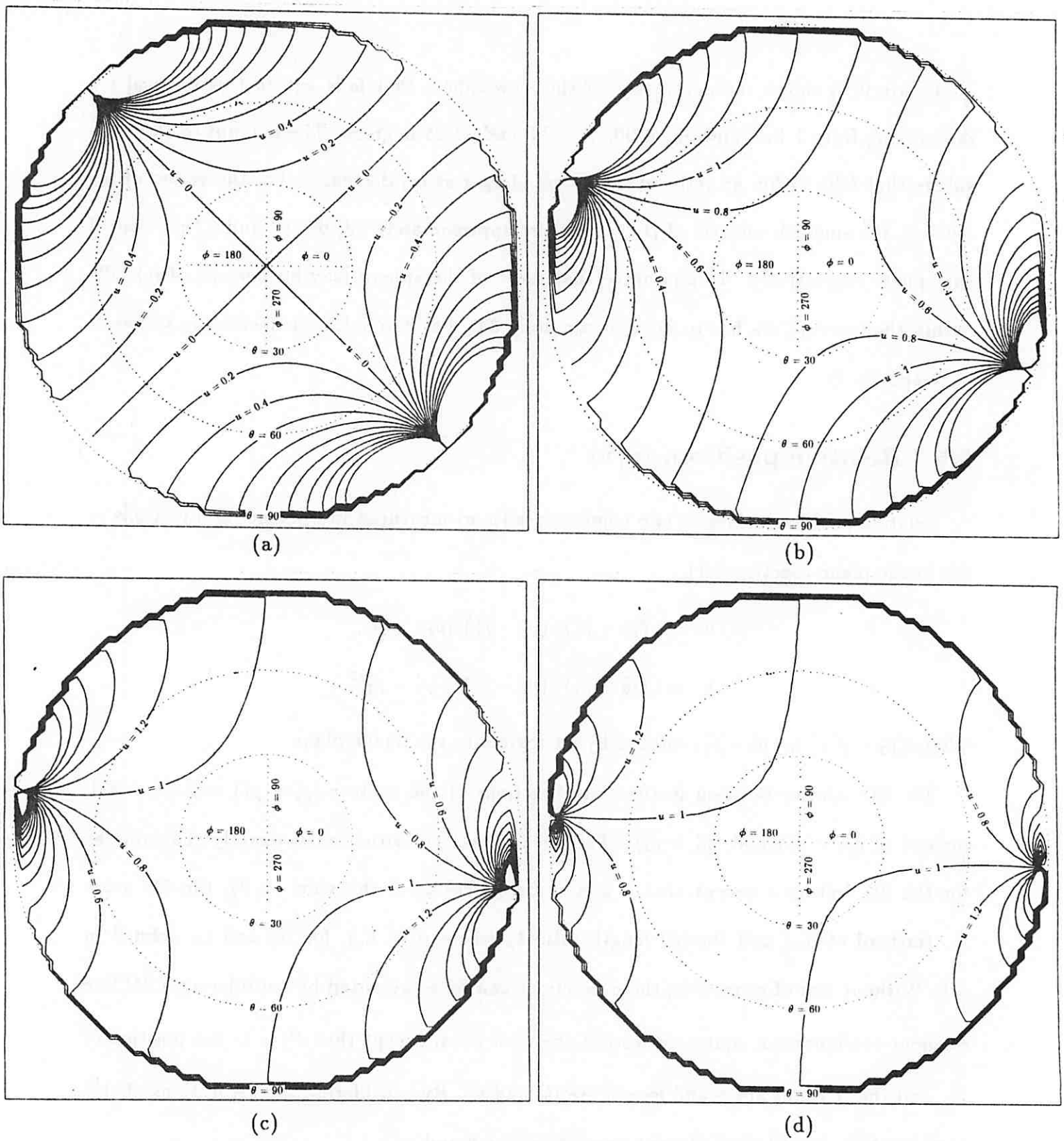


Figure 10: Plots of u as a function of θ and ϕ, for $t_w = 1$ and different values of β_w. The angle from each plot x axis represents ϕ (*azimuth*) and the radius from the center represents θ ($90 - elevation$). The contour lines indicate different values of u. The angle β_w (between the vectors $(\vec{p_2} - \vec{p_1})$ and $(\vec{p_3} - \vec{p_1})$) is (a) 90, (b) 45, (c) 22.5 and (d) 11.25 (degrees).

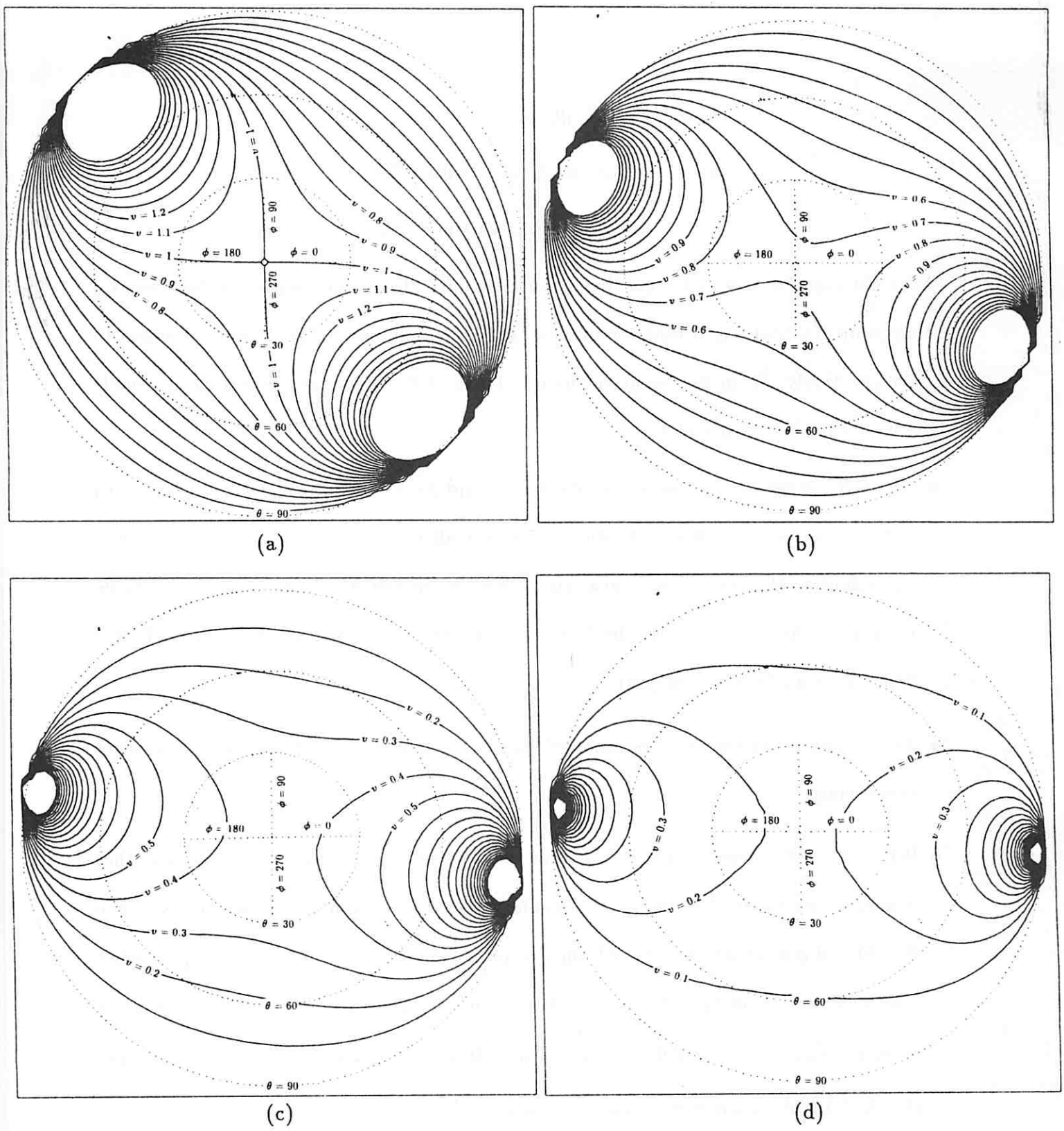


Figure 11: Plots of v as a function of θ and ϕ, for $t_m = 1$ and different values of β_w. The angle from each plot x axis represents ϕ (*azimuth*) and the radius from the center represents θ ($90 - elevation$). The contour lines indicate different values of v. The angle β_w between the vectors, $(\vec{p_2} - \vec{p_1})$ and $(\vec{p_3} - \vec{p_1})$, is (a) 90, (b) 45, (c) 22.5 and (d) 11.25 (degrees).

$$
\begin{aligned}
v &= \vec{p_3} \cdot \vec{p_2}^{\perp} / |\vec{p_2}|^2, \\
\vec{p_2} &= \pi R_{\phi,\theta}(\cos(\beta_w/2), -\sin(\beta_w/2), 0)^T \\
\vec{p_3} &= t_w \pi R_{\phi,\theta}(\cos(\beta_w/2), \sin(\beta_w/2), 0)^T
\end{aligned}
$$

As in the case for s, u and v are linear functions of a 3D distance ratio, in this case, t_w. Contour plots of u and v as a function of view (ϕ, θ), for different β_w, is shown in Figures 10 and 11 respectively. From the equations and plots, the following observations can be made for the view variation of u and v:

- When β_w is zero (i.e., the 3D points $\vec{P}_1$, $\vec{P}_2$ and $\vec{P}_3$ are collinear), u and v are both invariant to view, with $u = t_w$ and $v = 0$, for all t_w. In this 3D configuration, u is equivalent to the special-case invariant called the *approach ratio* in [Brooks81] (Table 1, entry 5), and v represents the fact that collinear sets of points project to collinear image points (Table 1, entry 6).

- As β_w approaches zero, the variation of u and v as a function of (θ, ϕ) is slow over most of the views.

- When $\beta_w > 0$, there always exists some view where $u = \infty$ and $u = -\infty$. In other words, u covers the full range of possible values. The value u approaches ∞ when $\theta = 90$ and ϕ approaches $-\beta_w/2$ from a counter-clockwise direction, and it approaches $-\infty$ when $\theta = 90$ and ϕ approaches $-\beta_w/2$ from a clockwise direction (the view is lining up with $(\vec{P}_2 - \vec{P}_1)$ from different directions). It approaches zero when (θ, ϕ) approach $(90, \beta_w/2)$ (this is when the view lines up with the displacement vector $(\vec{P}_3 - \vec{P}_1)$).

- When $\beta_w > 0$, there always exists some view where $v = \infty$ and $v = 0$ (v covers the full range of possible values). The value v approaches 0 when (θ, ϕ) approach $(90, \beta_w/2)$

and $(90, 180 + \beta_w/2)$. This is when the view lines up with $(\vec{P}_3 - \vec{P}_1)$). It approaches ∞ when (θ, ϕ) approach $(90, -\beta_w/2)$ and $(90, 180 - \beta_w/2)$ – when the view lines up with the first segment.

- As with the projected features already discussed, u and v are slowest in variation when $\theta = 0$ (i.e., the view is most normal to the plane containing the points $\vec{P}_1$, $\vec{P}_2$ and $\vec{P}_3$). At this point, $(u, v) = (t_w \cos\beta_w, t_w \sin\beta_w)$. Also, variation in u and v is still fairly slow for views near the Z axis, especially if the 3D angle β_w is reasonably small.

Figures 12(a) and 13(a) show the histograms of the number of views that fall within regular intervals of u and v, respectively, for various 3D angles β_w, with t_w set to one and the same view sampling scheme as in Section 5.3. The projected feature u actually ranges from minus infinitely to plus infinity, but for these examples, the histograms range over [-2, 2] and contain almost all of the samples. Similarly, v ranges from zero to infinity, but the range [0, 4] contains almost all of the samples.

As the histograms show, the distributions of the projected feature values are strongly concentrated about the value of the true 3D feature values $(u_w, v_w) = (t_w \cos\beta_w, t_w \sin\beta_w)$ and this concentration becomes more pronounced as β_w approaches zero (similarly for 180 degrees). A more revealing picture can be gained by counting the number of views for which u and v fall within intervals about the true feature values, $[u_w - \Delta u,\ u_w + \Delta u]$ and $[v_w - \Delta v,\ v_w + \Delta v]$.

Figure 12(b) shows the percentage of the view-sphere that falls within this interval for Δu ranging from 0 to 2 and β_w of 90, 45, 22.5 and 11.25 degrees. The amount of the view sphere that falls within an interval clearly gets larger as β_w decreases. For the values of β_w studied, a smallish interval of $[u_w - 0.2,\ u_w + 0.2]$ covers approximately 42, 52, 70 and 87

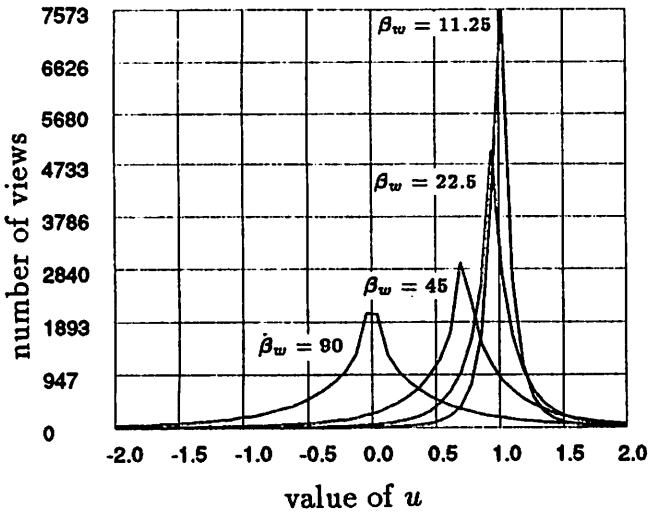


Fig. 12a Histograms of the number of views that fall within regular intervals of u for a hemisphere of views and for objects with various true 3D angles, β_w. (The range of u is divided into fifty intervals.)

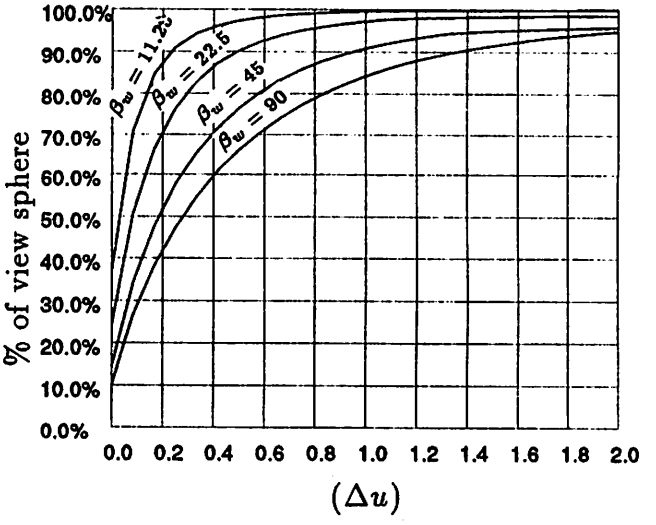


Fig. 12b The percentage of the view sphere that falls within the interval $[u_w + \Delta u,\ u_w - \Delta u]$ for various Δu ranging from 0 to 2 and β_w of 90, 45, 22.5 and 11.25 (degrees).

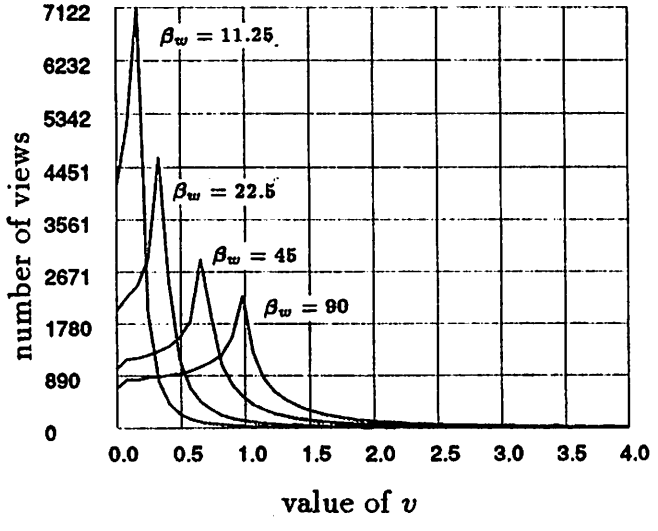


Fig. 13a Histograms of the number of views that fall within regular intervals of v for a hemisphere of views and for objects with various true 3D angles, β_w. (The range of v is divided into fifty intervals.)

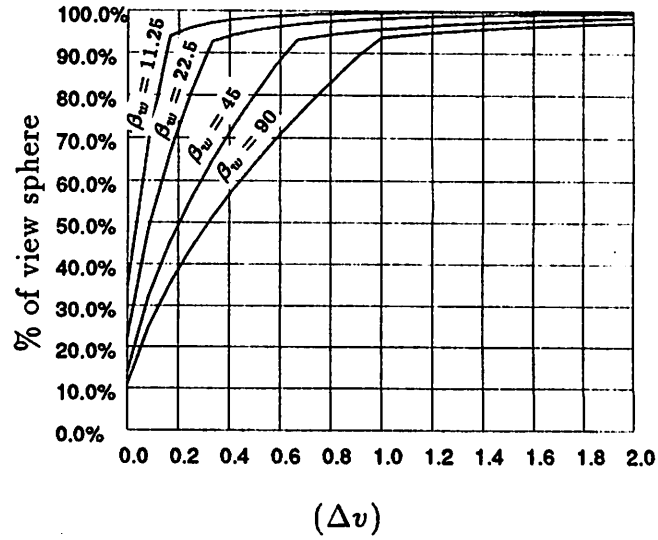


Fig. 13b The percentage of the view sphere that falls within the interval $[v_w + \Delta v,\ v_w - \Delta v]$ for various Δv ranging from 0 to 2 and β_w of 90, 45, 22.5 and 11.25 (degrees).

percent of the sphere respectively. To guarantee that most of the sphere (say 80 percent of it) falls within the interval, Δu has to be approximately .85, .58, .3 and .15 for each of the β_w respectively.

Figure 13(b) shows analogous graphs and behaviours for Δv ranging from 0 to 2. The smallish interval of $[v_w - 0.2,\ v_w + 0.2]$ covers approximately 38, 49, 73 and 95 percent of the view sphere for β_w of 90, 45, 22.5 and 11.25 degrees respectively. To guarantee that 80 percent of the views falls within the interval, Δv has to be approximately .75, .51, .25 and .12 for these β_w

6. Summary and conclusions

Though it has been shown here that there are no general-case view-invariants, there are special-case invariants of practical importance. For weak perspective, the special cases can be understood in terms of linear dependence and the invariance of this relation to linear transformation.

Also for weak perspective, there are 2D features with value distributions that tend to be restricted enough in the general case to provide useful features for 3D object recognition. The relative features of the projection, α, s, u and v are functions of two view parameters (θ, ϕ) and one or two 3D structure parameters (one for α and two for the others). All of these features vary slowest when θ is near zero and at this view assume the corresponding 3D features α_w, s_w, $u_w = t_w \cos\beta_w$ and $v_w = t_w \sin\beta_w$. When the directions or displacements are parallel ($\alpha_w = 0$ or $\beta_w = 0$) each feature becomes a special-case invariant, equivalent to some entry in Table 1. When they are non-parallel, each feature always varies across its whole range of possible values. In the case of s, u and v, the values "blow up" (i.e., assume

their extreme values) when the view aligns with a 3D displacement of importance to the feature. In the case of α, extreme values occur when the view lines up with the 3D angle bisector or a position in the object plane normal to it. In spite of the fact that the projected features vary across the full range of their values for most 3D line segment configurations, all of the features seem to vary reasonably slowly for a usable range of views: those views whose angle with the plane containing the object points is sufficiently large, but not necessarily near perpendicular. This is especially true when the 3D structure angles (α_w or β_w) are small. The observed concentration of the feature values about distribution peaks indicates the potential usefullness of the features studied and of prediction-based 3D recognition in general.

REFERENCES

[1] Ahuja, D.V. and S.A. Coons, "Geometry for Construction and Display", IBM Systems Journal, 7(3-4), pp. 188-205, 1968.

[2] Aliomonos, Y. and M. Swain, "Paraperpsective Projection: Between Orthography and Perspective", CAR-TR-320, U. Maryland, 1987.

[3] Binford, T.O., T.S. Levitt, W.B. Mann, "Baysian Inference in Model-Based Machine Vision", Proc. AAAI Uncert. Work., 1987.

[4] Brooks, R.A., "Symbolic Reasoning among 3D Models and 2D Images", AI, vol. 17, pp. 285-348, 1981.

[5] Burns, J.B. and L.J. Kitchen, "Recognition in 2D images of 3D Objects from Large Model Bases using Prediction Hierarchies", Proc. IJCAI-10, 1987.

[6] Duda, R. and P. Hart, *Pattern Classification and Scene Analysis*, Wiley-Interscience, 1973.

[7] Forsyth, D.A., J.L. Mundy, A. Zisserman, and C.M. Brown, "Projectively Invariant Representations Using Implicit Algebraic Curves", *Proc. First European Conf. Comput. Vision*, Springer LNCS, 1990.

[8] Grimson, W.E.L and T. Lozano-Perez, "Model-based Recognition and Localization from Sparse Range and Tactile Data", IJRR, vol 3, pp. 3-35, 1984.

[9] Ikeuchi, K. "Precompiling a Geometrical Model into an Interpretation Tree for Object Recognition in Bin-picking Tasks", DARPA IUW, pp. 321-339, 1987.

[10] Horn, B.K.P., "Robot Vision", MIT Press, 1986.

[11] Huttenlocher, D.P. and S. Ullman, "Object Recognition using Alignment", Proc. ICCV, pp. 102-111, 1987.

[12] Korn, M.R. and C.R. Dyer, "3D Multi-view Object Representations for Model-based Object Recognition", Pattern Recognition, vol. 20(1), pp. 91-103, 1987.

[13] Lamdan, Y., J.T. Schwartz and H.J. Wolfson, "Object Recognition by Affine Invariant Matching", Proc. CVPR, pp. 335-344, 1988.

[14] Lamdan, Y., H.J. Wolfson, "Geometric Hashing: A General and Efficient Model-Based Recognition Scheme", Proc. CVPR, pp. 238-249, 1988.

[15] Lowe, D.G., *Perceptual Organization and Visual Recognition*, Kluwer Academic Publishers, 1985.

[16] Marr, D. *Vision*, W.H. Freeman, 1982.

[17] Mohr, R., L. Morin, C. Inglebert and L. Quan, "Geometric Solutions to 3D Vision Problems", in preparation, 1990.

[18] Thompson, D.W. and J.L. Mundy, "Three-dimensional Model Matching from an Unconstrained Viewpoint", Proc. IEEE Int. Conf. on Rob. and Auto., pp. 208-220, 1987.

[19] Tucker, L.W., C.R. Feynman, D.M. Fritzsche, "Object Recognition using the Connection Machine", Proc. CVPR, 1988.

[20] Weiss, I., "Projective Invariants of Shapes", DARPA IUW, pp. 1125-1134, 1988.

A. Semi-regular sampling of spheres

For an accurate picture of the projected feature distributions as a function of view, it is important to sample the sphere of views as regularly as possible. Unfortunately, a truly regular sampling (one where the angular distance between all adjacent samples is exactly the same) is impossible beyond the 20 faces of a regular icosahedron [8].

A regular sampling of the spherical coordinates (ϕ, θ) (Section 5.2) produces a particularly non-uniform sampling of the sphere. This is precisely because the circles of constant θ have varying circumference: n regular samples of ϕ around circles of different θ result in a different angular distance between samples. Therefore, one way to produce an approximately regular sampling is by adjusting the number of samples n_θ around each circle θ. The number n_θ should be calculated so that the angular distance between adjacent samples on the circle is as close as possible to the angle between adjacent circles $(\Delta\theta)$.

[8] See [Horn86] p. 379 for further discussion.

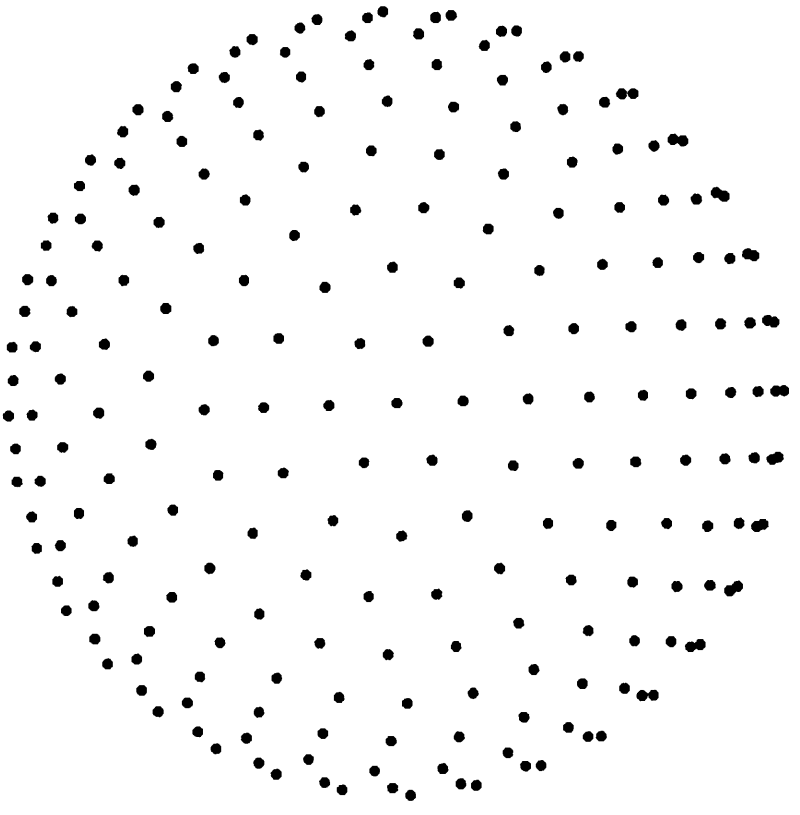


Figure 14: A semi-regular sampling of the sphere, viewed by orthographic projection. The angular distance between adjacent samples is approximately 10 degrees.

Assuming a unit sphere, the angular distance between two adjacent points, $\vec{P}_i$ and $\vec{P}_{i+1}$, on a circle at θ, is $\arccos(\vec{P}_i \cdot \vec{P}_{i+1})$. Positioning the first point arbitrarily at $\phi = 0$ gives:

$$\begin{aligned} \vec{P}_i &= (\sin\theta, 0, \cos\theta) \\ \vec{P}_{i+1} &= (\sin\theta\cos(2\pi/n_\theta), \sin\theta\sin(2\pi/n_\theta), \cos\theta) \\ \cos(\text{angular distance}) &= \sin^2\theta\cos(2\pi/n_\theta) + \cos^2\theta \end{aligned}$$

Setting the cosine of this angular distance to $\cos(\Delta\theta)$ and solving for n_θ results in:

$$n_\theta = \text{round}(2\pi/\arccos((\cos(\Delta\theta) - \cos^2\theta)/\sin^2\theta))$$

Figure 14 is an orthographic projection of a sphere sampling using this technique and a sample spacing of $\Delta\theta = 10$ degrees. (Note that the apparent shortening of the sample spacing near the image edge is due to projection.) When the sample spacing is set to one degree, the actual sampling has an average angular distance[9] of 0.9999 degrees with a standard deviation of 0.001 degrees.

[9]This is angular distance between adjacent samples at the same θ.