GROUPING CRITERIA IN ARCHITECTURAL SCENES FROM COMPATIBLE ORIENTATIONS OF POLYGONAL REGIONS

D.G. Aguilera[a], M. Gonzalo-Tasis[b], J. Finat[c]

[a] E. Politécnica Superior Ávila Univ. of Salamanca Spain daguilera@usal.es
[b c] MoBiVA Group, Univ. of Valladolid, 47011 Valladolid, Spain
[b] marga@infor.uva.es [c] jfinat@agt.uva.es

Commission V, WG V/4

ABSTRACT:

In this paper we address the problem of the estimation and automatic generation of structural perspective elements in architectural scenes. Structural character of geometric approach concerns to the preservation of incidence and adjacency constraints which can be visualized in different views of the same scene. The automatic generation of a perspective model from a view provides a support for the required visualization. It is necessary apply strong architectural constraints between the projection of structural elements lying onto planes of the 3D scene. These architectural constraints are relative to projections of parallelism and orthogonality conditions in the 3D scene dependent to elements lying onto at least two planes. We add grouping criteria based on quadrilaterals to include apparent distortions arising from perspective effects. Incidence relations between lines lying in adjacent planes in the 3D scene are visualized as junctions. The automatic identification of the type of junctions determines the relative orientation between adjacent planes. Adjacency and incidence properties allow us to construct a coarse piecewise linear structure based on oriented quadrilaterals. Their boundaries labelled as multiple junctions give a wired structure which is automatically superimposed to each view to simplify the generation of virtual views. The introduction of multivector representation simplify the visualization and management of 3D information arising from the lifting of adjacent quadrilaterals to 3D cuboids for visualization tasks.

1. INTRODUCTION.

Perspective models have been the main tool to represent 3D architectural scenes from the Renaissance. Increasing needs in Architecture, Photogrammetry, Motion Planning and Simulation require flexible and friendly interfaces which can be easily updated without human intervention. Then, it could be necessary generate automatically perspective models. which is based on the identification of pencils of perspective lines through vanishing points. Visualization of 3D information must identify pencils of planes through vanishing lines. Both of them require to identify *structural constraints* as colinearity and coplanariety, and *incidence constraints* for automatic grouping. Often, perspective lines in architectural scenes are manually drawn to avoid sensitivity errors. In our case, perspective lines are estimated by clustering of clouds of minisegments which are grouped following non-linear regression or RANSAC methods (to avoid the lack of robustness of least squares methods).

Depending of views, we can visualize inside the image, one, two or three vanishing points with their corresponding perspective models: frontal view, angular perspective, and oblique perspective, respectively. Selection of constraints in manually driven reconstruction depends on the perspective model chosen. In frontal and angular perspective , parallelism and perpendicularity constraints play a very important role, but they are less useful for oblique perspective. The increasing need for the integration of reconstruction and visualization tasks requires a more general approach to make compatible between them any perspective model in the projective framework.

Other problem correspond to data structures because depend on the chosen approach. There are essentially two approaches to recover and visualize a 3D structure : *features* and *primitive-based* approaches. In the current work, we have followed a primitive-based approach for frontal and angular perspective models.

In static scenes, mobile data structures correspond to egomotion or parallax. Our main contributions for perspective models with egomotion of the camera (including parallax effects) and their associated mobile data structures concern to a robust estimation of:
a) vanishing points,
b) parameters associated to transformations between 2D views,
c) their lifting to transformations of the ambient 3D projective space.

Robust and accurate estimation of vanishing points is obtained by adapting some well-known terrestrial photogrammetric methods. Parameters linked to transformations between views are expressed in terms of affine maps between quadrilaterals determined by intersections of perspective lines. Planar collineations between pairs of views are lifted to collineations in the ambient 3D space by constructing cuboids from the third vanishing point, as it is well known. Seemingly new aspect is relative to the symbolic management of information as propagation models from an initial cuboid. The robust identification of vanishing points allows to avoid cumbersome , "tedious" local compatibility and global coherence conditions for homologue cuboids in different constructions. It was indispensable restrict to simple indoor and outdoor scenes, avoiding occlusion problems and alternance between concave/convex regions, where monotonicity techniques arising from Computational Geometry can be applied.

The automatic management of convex quadrilaterals related to the intersection of pairs of pencils from perspective lines in each view is symbolically performed with quadtrees according to the quadrangular segmentation. Similarly the management of closed regions in the orthogonal 3D model M is performed in terms of octrees from triplets of pencils of perspective planes through three vanishing lines. External product and contraction in Geometric Algebra allow to transfer information between simplified 2D and 3D models.

The simplification of management of projective primitives and structural tools for simple 3D visualization, it is introduced from a symbolic representation given by quadtrees linked to templates of convex quadrilaterals in perspective models which are superimposed to each view. A convex quadrilateral is the image of a rectangle by an affine transformation in the plane. We lift quadtrees supported on planar templates given by perspective quadrilaterals, to octrees supported on volumetric templates given by cuboids. A cuboid is the image of a rectangle parallelepiped by an affine transformation in 3D space.

2. AUTOMATIC GROUPING TOOLS

Automatic grouping in Computational Geometry is usually performed by using auxiliary constructions such as decompositions in triangles, trapezoids or, more generally, convex sets associated to a multivectorized treatment of data. Minisegments are obtained by using a typical Canny's detector that, later, are grouped to determine ordinary multiple junctions. Multiple junctions are extracted by means of a variant of the Deriche's filter Character of these junctions informs about typical occlusions or about salient or entrant features.

The allowed types of multiple junctions are double, triple and quadruple (only allowed at vanishing points).. Typical junctions for a) two incident walls are L-type double junctions, b) inserted elements (doors, windows, etc) on a wall are T-type double junctions, c) corners associated to the perspective representation of trihedrals are Y- or arrow $\uparrow$ - type triple junctions, d) vertices of wireframed 3D representations or vanishing points linked to 4-tuples of lines are quadruple junctions. The automatic identification of collections of junctions along a closed polygonal allow us to generate facets, by including information about partially occluded regions, due to the relative localization of the camera. The comparison between regions in different views is reduced to find isomorphism between maximal ordered collections of junctions along candidate to be homologue polygonals.

An automatic interpretation of saliencies in the 3D scene is performed with corners labelled. To begin with, let us suppose that the sweep out has given us a segmentation by quadrilaterals supported on perspective lines. Each salient or entrant corner is characterized by three walls confluent at a typical Y-triple corner, i.e., a corner where discontinuities arising from extending visible segments are alternant with visible segments incident at such junction. So, in a typical architectural scene each triple junction Y is the common vertex for three quadrilaterals. The union of such three quadrilaterals give an hexagon H with the triple point inside connected to three vertexes of H which are also triple junctions.

The hexagon H inherits the natural orientation induced by compatible positively oriented quadrilaterals which are incident at the triple junction. If central triple junction is a right Y, then the central junction is salient. Otherwise, i.e. if central point is an inverted Y, then it corresponds to a corner of a concave region w.r.t. the observer. Typical hall indoor scenes or piecewise linear approach to baroque façades exhibit an alternant behaviour between right and inverse Y. The continuous or alternant character between triple junctions provide tools to connect local and global aspects. So, we obtain easily verifiable criteria for an automatic interpretation of the scene w.r.t. the observer's viewpoint.

Intersections of perspective lines determine a structure given by convex quadrilaterals which we overimpose to each image. For each pair of vanishing points $\mathbf{V_i}$ and $\mathbf{V_j}$ we generate a map of quadrilaterals; this map is specialized to a rectangular grid for orthogonal perspective and a trapezoidal grid for frontal view with a vanishing point in the view. The segments lying onto each quadrilateral of such a map make part of the pencil Λ_I and Λ_j of perspective lines through the vanishing points $\mathbf{V_i}$ and $\mathbf{V_j}$.

A coarse management of almost flat facades in outdoor scenes or flat walls in indoor scenes is performed with map of quadrilaterals. However, the sudden jumps in relative depth or the alternance between concave and convex regions (typical in Baroque style) reduce the performance of quadrilaterals maps. The solution of this problem involves the labelling of vertices onto the map of quadrilaterals as multiple junctions.

3. GENERATING PERSPECTIVE MODELS

Perspective models provide a prior knowledge of a 3D scene, which makes easier the interpretation, updating and insertion of new events. It is well known from the beginning of the XIXth century that pencils of lines and planes provide the general framework for *the management of perspective lines*. Thus, the first goal is to estimate pencils and their invariants. A pencil of lines $\mathbf{l}\ell + \mathbf{m}m$ is theoretically determined by two lines through the intersection point $\ell \cap m$. Nevertheless, the ratio of segments lying on two perspective lines is not invariant by perspective transformations. Consequently, it is convenient to take four lines in each pencil from at least four lines through a common point , eventually at infinity.

The intersection of two pencils Λ_I and Λ_j of perspective lines through two vanishing points $\mathbf{V_i}$ and $\mathbf{V_j}$ determine a map of quadrilaterals. In the same way as for trapezoids in frontal perspective in indoor scenes, we represent each quadrilateral by means of a *bivector* given as one half of the external product of the diagonal of the quadrilateral multiplied by the difference of vectors arising from a selected vertex of the diagonal. Thus, quadrilaterals inherit a natural orientation depending on the sweep out process and labelling processes. Each map of quadrilaterals generated by two planar pencil from perspective lines provides a support for the automatic grouping and for their associated bilinear maps. In fact, optimisation processes can be performed directly onto the space of bilinear maps, with the additional advantage linked to similarity relations between each map of quadrilaterals. In fact, propagation of similarities simplifies the matching process between candidate to homologue quadrilaterals. Above process can be performed for each pair of vanishing points. In this way, we obtain three maps of quadrilaterals which are matched together between them thanks to the existence of cuboids.

Every cuboid is algebraically represented by an oriented triple product. The transformation from cuboids to quadrilaterals it is done applying natural contraction operators in the framework of Geometric Algebra.

The automatic detection of alternance between concave and convex regions in the 3D scene is performed by using some criteria of Computational Geometry. Consequently, we develop a hierarchised strategy:

a) extraction of the upper and lower boundaries of regions located at the vertical planes in the scene,

b) segmentation of boundaries in the image depending on the relative orientation in the image,

c) evaluation of discontinuities in the orientation of the upper or lower boundary,

d) lifting of discontinuities to a rectified 3D model M,

e) generation of a map of oriented virtual planes

f) decomposition of our representation of the 3D scene in oriented volumes.

We illustrate this strategy with examples with superimposed automatically and manually generated structures. It is necessary avoid the expensive tests for 3D visibility by a beforehand evaluation of reflexivity chains for the boundaries in the image which is based on monotonicity properties.

4. ROBUSTNESS AND ACCURACY

In traditional photogrammetry common features are matched in a very accurate way by hand. The operator's cost and the increasing needs and recent advances in Computer Vision have impulsed the development of automatic matching tools. Vanishing points and lines provide structural elements for perspective models which are robust w.r.t. small perturbations. Hence, an accurate identification of vanishing points allows to retrace perspective lines through vanishing points, and to superimpose perspective planes through vanishing lines. Retracing of lines reinforces colinearity constraints, including dual information relative to pencils of lines through vanishing points. In the same way, the construction of perspective planes through vanishing lines reinforces coplanariety constraints to visualize the 3D scene.

The first author has adapted several robust estimators for vanishing points from perspective lines arising from terrestrial photogrammetry (danish method, minimal sum, Huber, German -McClure). The main goal of all of them is the automatic elimination of lines which are not adequate for the "right" determination of vanishing points. The rightness is measured in terms of the area minimization. In typical architectural scenes we have a high redundancy degree, more precisely, we have enough elements arising from an automatic vectorization to provide bundle of errors in perspective lines. In this case, this method is revealed accurate and robust.

However, in some indoor scenes, due to bad illumination problems, mainly or in outdoor scenes with few architectural elements: traffic in highway, e.g., sometimes it is difficult to identify perspective elements to visualize 3D scenes. In this case, we adopt a coarser approach based on a feedback between putative perspective lines (by regression and Hough transform), determination of vanishing points by vote and importance schemes [Tor03] and retracing of lines. Results obtained by second and third authors are less accurate, but they fulfil some standard requirements such as robustness and real-time processing.

To integrate both approaches in a common optimization framework, it is necessary to formulate the best criteria, and the right formulation depends on the sources of unstability.

Obviously, there is no a unique unstability source, because it depends on the scene and the different kinds of motion allowed in the scene. The space of multilinear maps defined on a projective model for the ambient space provides a general framework.

Optimization on the space of multilinear maps must include some kind of feedback based on the coherence of 2D templates of quadrilaterals for the analysis of planar views and 3D templates of cuboids for their visualization The key fact is that planar convex quadrilaterals provide a geometric support for bilinear maps, whereas cuboids provide similar geometric support for trilinear maps. It is crucial to avoid the deterioration of accuracy linked to occasional features, by the estimation of invariant data measured on bundles of incident lines (for planar models) through a vanishing point or incident planes (for volumetric visualizations) through a vanishing line. In both cases, the cross-ratio of four elements of pencil is the main invariant, essentially unique for the linear case, in fact.

5. SYMBOLIC MANAGEMENT OF ARCHITECTURAL DATA STRUCTURES

The best photorealistic rendering of a 3D scene corresponds to a dense matching of pixels. However, the high computational cost, and the need for updating in presence of egomotion suggests to use a hybrid approach able of combining features and primitives in a common geometric framework.

To avoid a cumbersome dense information, with the corresponding high computational cost, we follow a multilayered interpretation and reduce ourselves to data contained in perspective models which are superimposed to the original views.

Doubly and triply connected lists linked to quadrilateral and cuboid superimposed structures provide tools to simplify data structures relative to perspective models of planar views and their lifting to 3D visualization, respectively.

The symbolic management of this lifting is performed by means a simplified version of the blow-up of graph with center a triple corner linked to a vanishing point. In the simplest case, the blowing-up of a planar polygonal oriented facet F w.r.t. a point away from the plane, replaces such point by an oriented copy of the F located in a parallel plane.
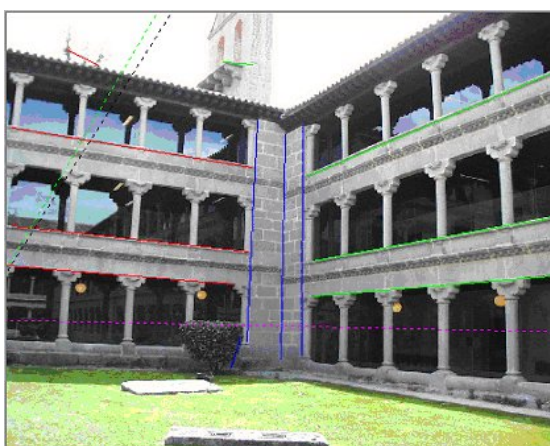
Then it is possible to obtain a coarse information about occupied or empty regions from the observer's viewpoint It must be understood that blowing-up processes relative to triple corners is strictly local ,always w.r.t. to opposite sides in perspective models. However, the global interpretation in terms of direct or inverse type Y triple junctions, provides an easy interpretation of matching processes.

6. EXPERIMENTAL RESULTS

We have applied some robust estimators to several images obtained of Santa Ana's Cloister. The results of perspective lines and vanishing points can be showed in the following images:
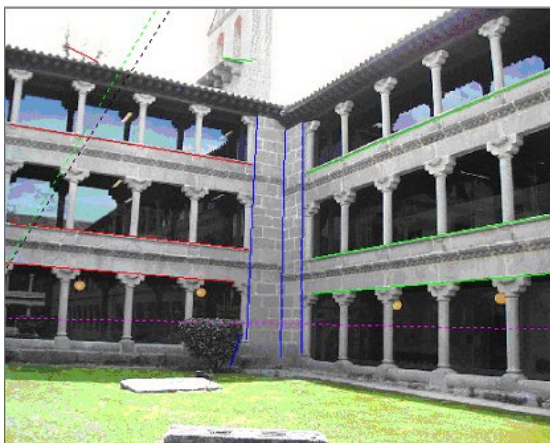
(a) Vanishing Points after applied danish estimator

(a) Vanishing Points after applied Classical Estimator LSM

7. CONCLUSIONS AND FUTURE WORK

We have developed several methods for a robust estimation of vanishing points inside the views corresponding to projections of indoor and outdoor scenes. Automatic retracing of perspective lines through vanishing points provides a perspective template given by a map of quadrilaterals. The map of quadrilaterals condensates the grouping criteria with their structural constraints for each view. Similarly, we construct a contraction of octrees to quadtrees in terms of blowing-down graphs depending on visibility constraints.

Visibility constraints are directly linked to the estimation of oriented facets and, consequently with the vanishing of volume forms linked to volumetric elements in the 3D representation of the scene.

REFERENCES

[Fin02] J.Finat , M.Gonzalo, M.J. Antolinez y S. Aguilar: "Dynamic Trapezoidal Maps For Coarse Perspective Models In Indoor Scenes", ISPRS'02

[Kan93] K.Kanatani: "Geometric Computation for Machine Vision", Oxford Univ. Press, 1993.

[Tor03] P.H.S.Torr and C.Davidson: "IMPSAC: Synthesis of Importance Sampling and Random Sample Consensus", IEEE Trans on PAMI, Vol.25, nº3, 2003, 354-364.