DIPLOMATERVEZÉSI FELADAT

Danyi Dávid

Villamosmérnök hallgató részére

Marker alapú helymeghatározás képfeldolgozással

Az információfeldolgozási kapacitás nagymértékű növekedésével párhuzamosan egyre szélesebb körben váltak alkalmazhatóvá (valós időben) képfeldolgozási, gépi látás alapú eljárások. Az ideálistól eltérően a valós képek feldolgozását számos hatás nehezíti, úgy mint zaj, optikai torzítások, megvilágítás- és színbeli különbségek, stb.

A feladat témája aktuális, az egyre gyakrabban alkalmazott mobil robotikai, kiterjesztett és virtuális valóság alkalmazásoknál minden esetben felmerül a „hol vagyok?" kérdés, azaz a nézponti paraméterek becslése, lokalizáció. A SLAM (Simultaneous Localization and Mapping) algoritmus számos szenzorforrás által biztosított információval alkalmazható, így kamerával is. Az általános megoldás tájékozódási pontok egymáshoz képesti elhelyezkedését becsli, a mérési bizonytalanságot folyamatosan csökkentve, valamint méri a tájékozódási pontokhoz képest a megfigyelő pozícióját.

Jelen munka is a fenti témához, kapcsolódik, cél megvizsgálni egy olyan mesterséges, passzív marker megvalósíthatóságát, ami bizonyos paraméterekben (azonosíthatósági tartomány, részleges láthatóság) jobbat kíván biztosítani, mint az elterjedt (ARtag, glyph, stb.) megoldások. A marker egy oldalukon nyitott különböző méretű és alakú négyszögeket tartalmaz. A diplomaterv célja megvizsgálni a javasolt marker jellemzőit és használhatóságát.

A hallgató feladatának a következőkre kell kiterjednie:
- Mutasson be pozícióbecslő algoritmusokat, röviden ismertesse ezek működését!
- Hasonlítson össze különböző nézőpontbecslési algoritmust síkban elhelyezkedő pontpárok alapján!
- Készítsen egy marker felismerő megoldást!
- Végezzen méréseket (ideális és valós képeken) a marker által meghatározott pozíció pontosságára vonatkozóan!
- Hasonlítsa össze és értékelje az eredményeket!

Tanszéki konzulens: Kovács Viktor, tanársegéd

Budapest, 2017. február 18.

Dr. Charaf Hassan
egyetemi tanár
tanszékvezető

Budapest Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Automatizálási és Alkalmazott Informatikai Tanszék

H-1117 Budapest, Magyar tudósok krt 2. Q.B.207
Telefon: 463-2870. Fax: 463-2871
www.aut.bme.hu

MŰEGYETEM 1782

Budapest University of Technology and Economics

Faculty of Electrical Engineering and Informatics

Department of Automation and Applied Informatics

Marker Based Localisation and Pose Estimation Using Image Processing

THESIS

Author	*Consultant*
Dávid Danyi	Viktor Kovács

May 11, 2017

Contents

HALLGATÓI NYILATKOZAT

Alulírott *Dávid Danyi*, szigorló hallgató kijelentem, hogy ezt a szakdolgozatot meg nem engedett segítség nélkül, saját magam készítettem, csak a megadott forrásokat (szakirodalom, eszközök stb.) használtam fel. Minden olyan részt, melyet szó szerint, vagy azonos értelemben, de átfogalmazva más forrásból átvettem, egyértelmûen, a forrás megadásával megjelöltem.

Hozzájárulok, hogy a jelen munkám alapadatait (szerzõ(k), cím, angol és magyar nyelvû tartalmi kivonat, készítés éve, konzulens(ek) neve) a BME VIK nyilvánosan hozzáférhetõ elektronikus formában, a munka teljes szövegét pedig az egyetem belsõ hálózatán keresztül (vagy autentikált felhasználók számára) közzétegye. Kijelentem, hogy a benyújtott munka és annak elektronikus verziója megegyezik. Dékáni engedéllyel titkosított diplomatervek esetén a dolgozat szövege csak 3 év eltelte után válik hozzáférhetõvé.

Budapest, May 11, 2017

_____

Dávid Danyi
hallgató

Abbreviations

This is a complete list of the abbreviations used in this paper.

DOF Degrees of freedom

RQIM Random Quad Image Marker

Chapter 1

Markers

One of the goals of this project was to design a fiducial marker with advantageous properties for use in pose estimation. In a typical scenario the marker may be seen from largely varying viewpoints. Therefore it has to have some level of scale invariability. If the observer is far from the marker, the smaller details may be lost due to the limited resolution of the camera. If the same observer moves closer to the marker, it may fill the whole field of view and some features may even slip off the image. This leads to another feature the marker needs: redundancy. If the observer gets too close to the marker or some obstacle partially blocks the view, the localisation still needs to provide usable results.

The intended use of the markers is spatial localisation and pose estimation. In other words: approximating the observers 3D coordinates (x, y, z) and orientation (ϕ, θ, ψ) with respect to the marker. It is supposed that the observer uses a single camera system for navigation (e.g. smartphone or robotic application with limited resources). This means the marker needs at least 6 degree of freedom.

To sum up the above discussed specifications, a suitable marker would have to:

- have at least 6 DOF

- be (to some degree) scale invariant

- have redundancy

In the following sections will be a recommendation for a marker conforming for the listed specifications. It is based on 3 connected line segments forming a quad with one missing side. The whole marker is built from quads with different sides and angles.

1.1 Quad

A marker is put together from quads. Figure 1.1. shows two examples. One side of the quads is left out: they are put together from three joint line segments. The middle segment, with
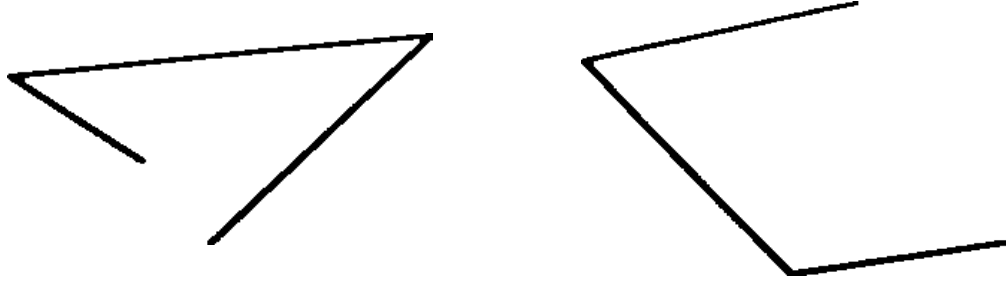
Figure 1.1: *Example for different quads*

two adjoining lines, will be referred to as the 'base' of the quad. The outer segments are going to be called 'arms'.

A quad has 6 degrees of freedom. There are 3 independent distance parameters: the length of the base and the two arm segments. There are also 3 unrelated angle parameters: the angles between each arm and the base, and the orientation of the quad.
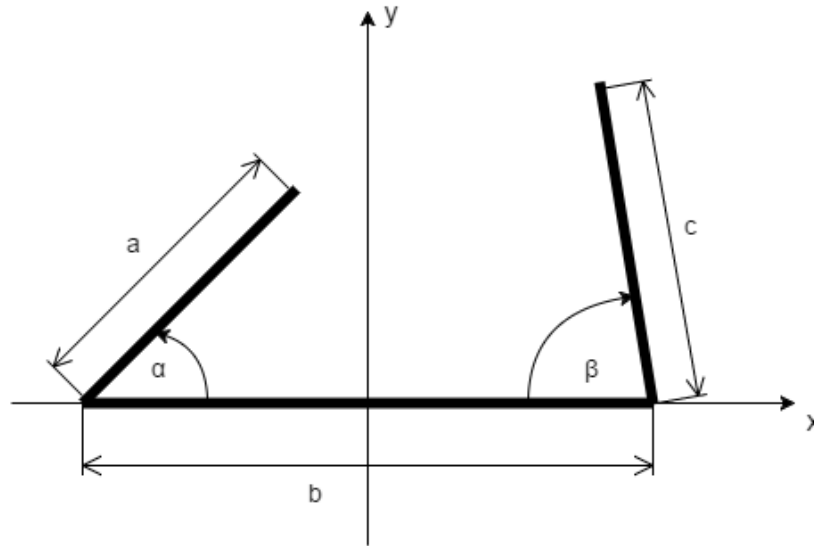

Figure 1.2: *Quad parameters*

Figure 1.2. shows the free parameters of a quad (the orientation is not shown on the image). The following notation is used:

- a : The length of one arm

- b : The length of the base

- c : The length of the other arm

- α : The angle between one arm and the base

- β : The angle between the other arm and the base

- γ : The angle with which the whole quad is rotated

For the sake of simplicity, figure 1.2. does not show the rotation with γ. The quad would be rotated around the origin of it's coordinate system.

The value of the length parameters are given in pixels, although they can be expressed in any unit of distance. The angles can be given is degrees of radians (in the implementation degrees are used for easier human readability).

$$a \in (0, a_{max}] \tag{1.1}$$

$$b \in (0, b_{max}] \tag{1.2}$$

$$c \in (0, c_{max}] \tag{1.3}$$

$$\alpha \in (0, 180°) \tag{1.4}$$

$$\beta \in (0, 180°) \tag{1.5}$$

$$\gamma \in [0, 360°) \tag{1.6}$$

Equations (1.1) through (1.6) specify the range of each parameter. The maximum of the distance parameters are set by the space left on the image for the given marker, there is no theoretical limit for them. There is also no constraint for the resolution of the parameters. From the applications point of view, there are quads with continuous[1] and discrete parameter spaces.

1.1.1 Quad representation

There are several ways to represent quads, each with different advantageous properties. For this work multiple considerations were made in that regard. The most straightforward is to simply store the above mentioned parameters. This is simple and easy for human reading, which is great help in the development process.

A step forward from this is to norm the a, b and c parameter of the quad with the base segment's length. Then the following parameters are used:

s : marker size, the same as the base length

m_a : 'a' multiplier. $m_a = a/b$

m_c : 'c' multiplier. $m_c = c/b$

The α, β, γ angle parameters are not changed. This gives a scale or size parameter for the quad, which is useful for marker generation. These two representations are good for development and marker generation, but not so much for calculations.

---

[1]That is, only limited by the computational resolution

A third option is to store the endpoints of the line segments. It requires the storage of 4 points: two endpoints (E_1, E_2) and two inner points (I_1, I_2). Equations (1.7) through (1.10) define the points' coordinates before rotating with γ.

$$I'_1 = (-\frac{b}{2}, 0) \tag{1.7}$$

$$I'_2 = (\frac{b}{2}, 0) \tag{1.8}$$

$$E'_1 = (-\frac{b}{2} + b * cos(\alpha), a * sin(\alpha)) \tag{1.9}$$

$$E'_2 = (\frac{b}{2} - c * cos(\beta), c * sin(\beta)) \tag{1.10}$$

$$Rot(\gamma) = \begin{pmatrix} cos(\gamma) & -sin(\gamma) \\ sin(\gamma) & cos(\gamma) \end{pmatrix} \tag{1.11}$$

The point E_1, E_2, I_1, I_2 can be obtained from E'_1, E'_2, I'_1, I'_2 by a multiplication with the rotational matrix $Rot(\gamma)$.

That method is redundant for storage: it uses 8 parameters instead of 6. However this poses no practical problem in the scope of the project. The one outstanding benefit of this method is it's efficiency in calculations. Because it is based on points in a linear space, linear algebraic methods (matrix multiplications) can be used for calculating he projective transformations.

In this project the second and the third options are used. The first, naive method is omitted because it has no considerable advantage over the other two. The second method, using the size, multiplier and angle parameters is used in marker generation. The third is used during the calculations and the recognition process.

1.2   Marker

Quads are 6 DOF shapes: in theory it would be enough to use only one of them for localisation and pose estimation. However that method would have very low error tolerance and questionable accuracy even in a best case scenario. To comply with the specifications written in the beginning of this chapter, the markers are put together from multiple quads. By placing quads with different orientations and sizes the error tolerance and accuracy can be greatly improved.

An intrinsic positive quality of using multiple quads with varying sizes is the scale invariance. As mentioned, even a single quad is sufficient for the task at hand. If the smaller quads become unrecognisable because of the low resolution or too large distance, a successful measurement is still possible. The same is true on the other end of the spectrum: if the observer is too close to the marker and the larger ones leave the field of view, the position and orientation can be calculated from the smaller quads.

Figure 1.3. shows an example for a marker. It is generated with naive algorithm described in the next section, and is not optimal in many ways. Nonetheless it is functional, even if only a fraction of the quads are registered for the measurement.
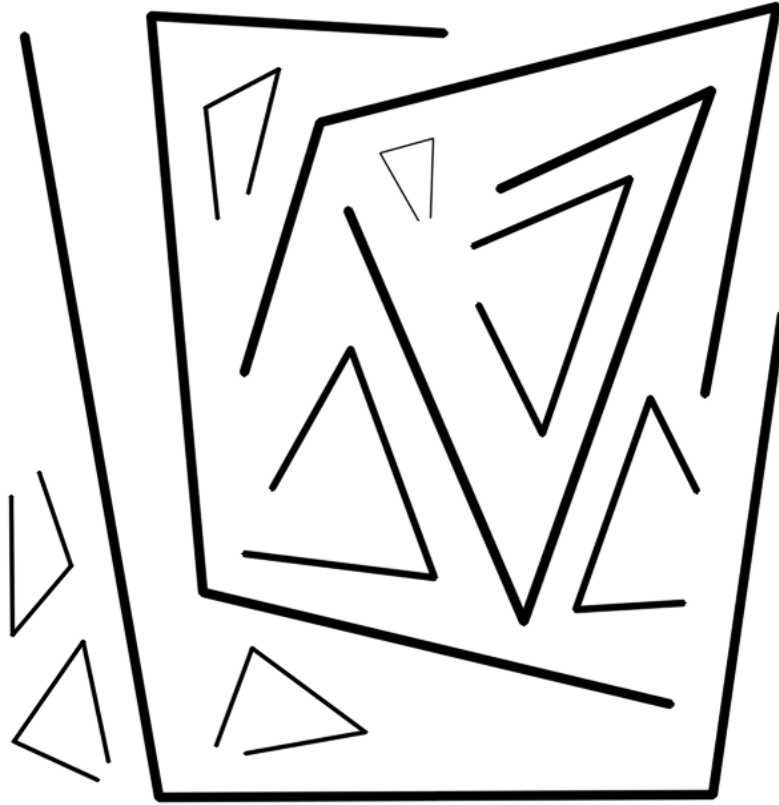

Figure 1.3: *An example for a marker*

The markers are going to be referred to as RQIM, which means Random Quad Image Marker. As the name suggests, the quads are randomly generated and placed on the markers.

Unless otherwise specified, RQIMs use quads with continuous parameter spaces. In this chapter there will also be a small introduction to the discrete parameter markers and their potential applications.

1.2.1 Marker generation

1.2.2 Discrete RQIM

Chapter 2

Marker recognition

In this chapter there will be a summary of the image processing algorithms tried and used for the recognition of the fiducial markers. From a computer vision point of view the task is to detect joint line segments.