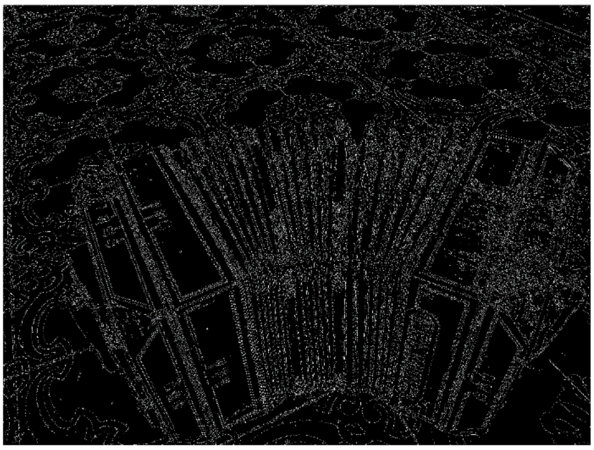
Image Analysis and Computer Vision
Year: 2017/18
Professor: Vincenzo Caglioti

Homework

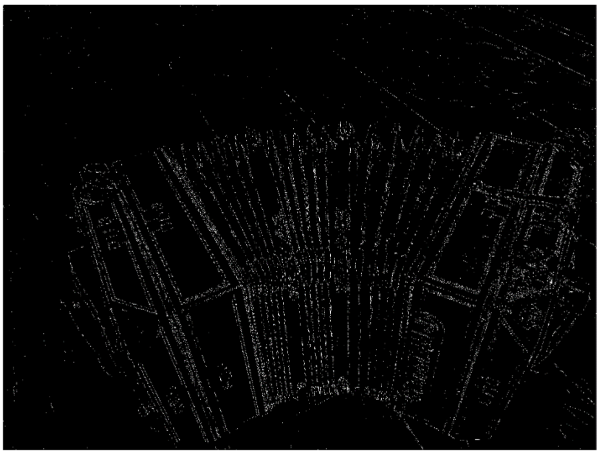
Matheus Henrique Fim – 876069

1. Image Feature Extraction and Selection

In order to find the edges in the image it was made a comparison between the Canny, Sobel and Prewitt detection methods, from which Canny yielded clearer results, detecting also the patterns in the floor.
In the Binary image outputted by the edge function, it was applied the Hough transform, that allowed to plot the lines in the image.

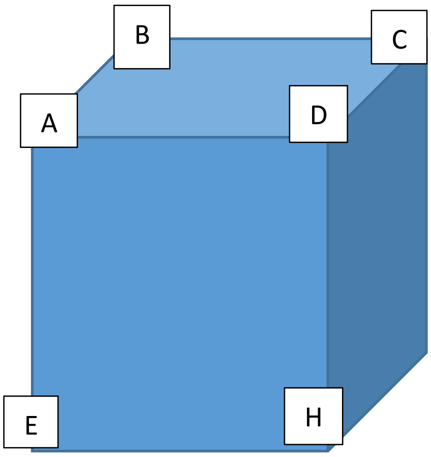


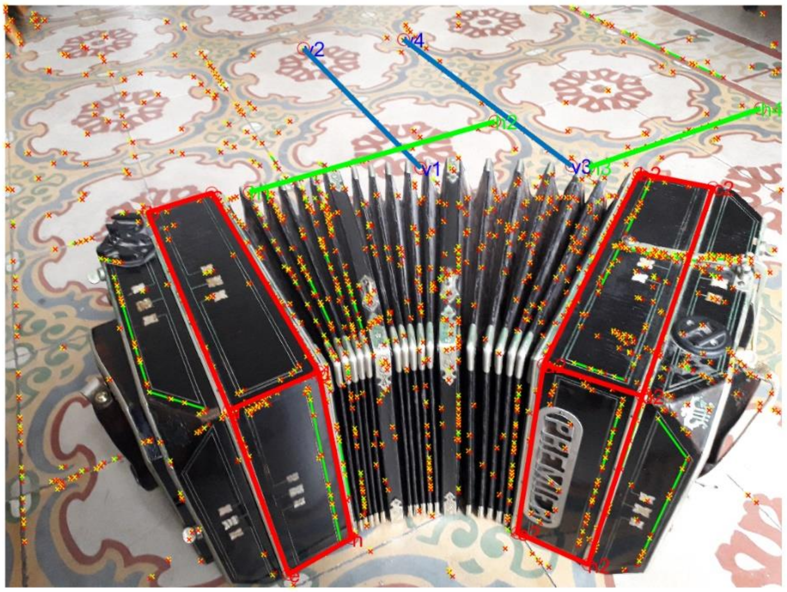
Canny



Sobel

To conclude the first step, the lines that are useful to the subsequent tasks were individuated through the "getpts" function. Those lines were two parallel horizontal lines on the floor, two parallel vertical lines on the floor and the points corresponding to the a, b, c, d, e, h position in each of the bandoneon wooden parts.
The convention adopted when naming the cube was as follows in the picture bellow





2. Geometry

The first task was to apply 2d rectification and thus reconstruct the shape of the bandoneon horizontal faces. This task was further divided in two steps: remove projective distortion and remove affine distortion.

2.1 Remove Projective Distortion

One of the ways of removing the projective distortion is to calculate the vanishing line and find the homography that maps it back to $l\infty$.

A simple way of doing this is to find two line pairs (L1' ,L2') and (M1', M2') in the image that are supposed to form two different parallel lines in the world plane, calculate their vanishing points:

$$Vp1 = L1' \times L2'$$

$$Vp2 = M1' \times M2'$$

And from here derive the vanishing line Vl:

$$Vl' = Vp1 \times Vp2$$

Consequently the homography that we are looking for is given by:

$$HP = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ Vl_1 & Vl_2 & Vl_3 \end{vmatrix}$$

One important detail is that this algorithm is heavily affected by the vanishing line precision, and thus we ought to determine it in the most accurate possible way. Keeping this in mind, besides the lines in the bandoneon horizontal faces, it was also used the lines in the floor. The strategy was to calculate the vanishing point of each two parallel lines and then use the algorithm described by Bob Collin to fit the best possible vanishing line.

2.1.1 Bob Collin Best fit algorithm:

Considering the set of lines l1, l2.. ln.

- form the matrix

$$M = sum \begin{vmatrix} a\_i * a\_i & a\_i * b\_i & a\_i * c\_i \\ a\_i * b\_i & b\_i * b\_i & b\_i * c\_i \\ a\_i * c\_i & b\_i * c\_i & c\_i * c\_i \end{vmatrix}$$

where the sum is taken for each of the n lines.

- Perform an eigen decomposition of M, and select the smallest eigen value as the vanishing point.

2.2 Remove affine distortion

After correcting for the projective distortion one can notice that the real world angles are not preserver in the image.

In order to correct it we only need two pairs of mutually orthogonal lines in the plane. This is due that the angle between the orthogonal lines is given by:

$$\cos\theta = \frac{L^T C_\infty^* M}{\sqrt{(L^T C_\infty^* L)(M^T C_\infty^* M)}}, C_\infty^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Amd since cos θ = 0, we can derive that:

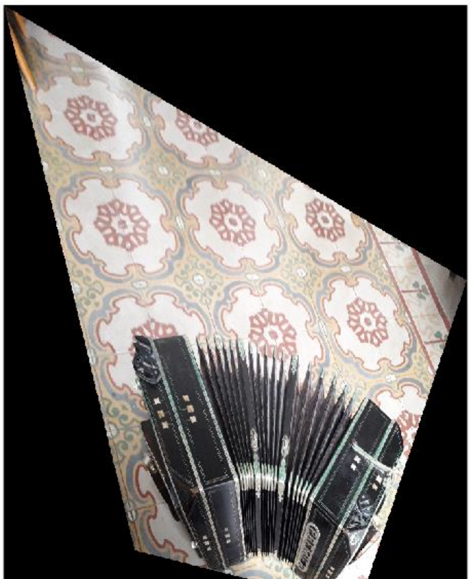
$$L'^T H_a C_\infty^* H_a^T M' = 0$$

Where Ha is the homography that removes the affine distortion.
Expanding the above equation, we arrive to:

$$[l_1' \quad l_2' \quad l_3'] \begin{bmatrix} AA^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} m_1' \\ m_2' \\ m_3' \end{bmatrix} = 0$$

Where S = transpose(AA) = [s11, s12 ; s12, s22]

There are only two degrees of freedom, since it is symmetrical and S22 can be set to 1 as it only the ratios are important. Furthermore, using SVD on S we can arrive to A, and thus, to Ha.



Rectified image

2.3 Calibrating the camera

To determine the intrinsic parameters of the camera we are going to make use of the image of the absolute conic, that is given by

$$\boldsymbol{\omega} = \begin{bmatrix} w_1 & w_2 & w_4 \\ w_2 & w_3 & w_5 \\ w_4 & w_5 & w_6 \end{bmatrix}$$

We already know the the camera has zero-skew, this w2 = 0, leaving us with four non trivial elements. In other words, we need to set other 4 equations.

Condition	constraint	type	# constraints
vanishing points $\mathbf{v}_1, \mathbf{v}_2$ corresponding to orthogonal lines	$\mathbf{v}_1^{\mathsf{T}} \omega \mathbf{v}_2 = 0$	linear	1
vanishing point $\mathbf{v}$ and vanishing line $\mathbf{l}$ corresponding to orthogonal line and plane	$[\mathbf{l}]_\times \omega \mathbf{v} = \mathbf{0}$	linear	2
metric plane imaged with known homography $\mathtt{H} = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$	$\mathbf{h}_1^{\mathsf{T}} \omega \mathbf{h}_2 = 0$ $\mathbf{h}_1^{\mathsf{T}} \omega \mathbf{h}_1 = \mathbf{h}_2^{\mathsf{T}} \omega \mathbf{h}_2$	linear	2
zero skew	$\omega_{12} = \omega_{21} = 0$	linear	1
square pixels	$\omega_{12} = \omega_{21} = 0$ $\omega_{11} = \omega_{22}$	linear	2

Table 8.1. **Scene and internal constraints on ω.**

In accordance with the above table, extracted from the text book, the metric plane imaged with known homography provides 2 constraints. While the relation between the vanishing point of the vertical face of the bandoneon with the vanishing line previously discovered will give us two more equations. From this point it W can be straightforwardly determined by the SVD function. While Cholesky decomposition will give the intrinsic parameters matrix K.
The extrinsic parameters will be given by the equations:

$\boldsymbol{h}_0 \leftarrow \mathbf{H}_{*,0}$ $\qquad \triangleright\ \boldsymbol{h}_0 = (H_{0,0}, H_{1,0}, H_{2,0})$
$\boldsymbol{h}_1 \leftarrow \mathbf{H}_{*,1}$ $\qquad \triangleright\ \boldsymbol{h}_1 = (H_{0,1}, H_{1,1}, H_{2,1})$
$\boldsymbol{h}_2 \leftarrow \mathbf{H}_{*,2}$ $\qquad \triangleright\ \boldsymbol{h}_2 = (H_{0,2}, H_{1,2}, H_{2,2})$
$\kappa \leftarrow 1/\|\mathbf{A}^{-1} \cdot \boldsymbol{h}_0\|$
$\mathbf{r}_0 \leftarrow \kappa \cdot \mathbf{A}^{-1} \cdot \boldsymbol{h}_0$
$\mathbf{r}_1 \leftarrow \kappa \cdot \mathbf{A}^{-1} \cdot \boldsymbol{h}_1$
$\mathbf{r}_2 \leftarrow \mathbf{r}_0 \times \mathbf{r}_1$ $\qquad \triangleright$ 3D cross (vector) product
$\boldsymbol{t} \leftarrow \kappa \cdot \mathbf{A}^{-1} \cdot \boldsymbol{h}_2$ $\qquad \triangleright$ translation vector
$\tilde{\mathbf{R}} \leftarrow (\mathbf{r}_0 \mid \mathbf{r}_1 \mid \mathbf{r}_2)$ $\qquad \triangleright$ initial rotation matrix

While R may be transformed in a true rotation matrix by applying SVD function and assigning R=U*V'.

2.4 Camera Pose

We can then calculate the Projection matrix

P = K*[R', t]

That is also noted as:

$$\mathtt{P} = [\mathbf{l}_1^{\perp} \quad \mathbf{l}_2^{\perp} \quad \alpha\mathbf{v} \quad \hat{\mathbf{l}}] \qquad (1)$$

Which is then used to discover the camera positions

$$
\begin{aligned}
X_c &= -\det\begin{bmatrix}\mathbf{l}_2^{\perp} & \mathbf{v} & \hat{\mathbf{l}}\end{bmatrix}, \quad Y_c = \det\begin{bmatrix}\mathbf{l}_1^{\perp} & \mathbf{v} & \hat{\mathbf{l}}\end{bmatrix} \\
\alpha Z_c &= -\det\begin{bmatrix}\mathbf{l}_1^{\perp} & \mathbf{l}_2^{\perp} & \hat{\mathbf{l}}\end{bmatrix}, \quad W_c = \det\begin{bmatrix}\mathbf{l}_1^{\perp} & \mathbf{l}_2^{\perp} & \mathbf{v}\end{bmatrix}
\end{aligned}
$$

Up to a scale factor alpha, that is yet unknown.

2.5 Scale Factor

To make possible to determine the homographies of different planes we still need to determine the scale factor of each column in our projection matrix. This can be achieved by using the information on the size of the long side of the horizontal face, from there we can determine the scale factor of Vy. With this scale factor it is possible to measure the size of the short side on the horizontal face, which happens to be common to a vertical face as well. Thus we can derive the scale factor for Vx, and finally we arrive at the scale factor of Vz. The formula used is:

$$
\alpha Z = \frac{-||\mathbf{b} \times \mathbf{t}||}{(\hat{\mathbf{l}} \cdot \mathbf{b})||\mathbf{v} \times \mathbf{t}||}
$$

Where alpha is the scale factor, Z is the measure; B and T are a base point in the reference plane and a top point in the scene, respectively.

$$
\mathbf{b} = \mathtt{P}\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{t} = \mathtt{P}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}
$$

2.6 Homographies

At last we can apply the scale factors to the projection matrix and arrive to the homography of the vertical face of the bandoneon.