Moment Invariant Shape Features: a brief explanation

Jonathan Campbell, Department of Computing, Letterkenny Institute of Technology, Co. Donegal, Ireland, jonathan.campbell@lyit.ie. 23 March 2004.

Introduction

Invariant moments were introduced by (Hu 1962); elaborations and discussions of computer implementations are discussed in (Masters 1995).

What we require are features that capture the *shape* of objects. Since an object can appear in any position in the image, we require *shift invariant* features; likewise, since an object must be recognised independently of its angular orientation, we require *rotationally invariant* features. Finally, because we want to remove image magnification and other size effects, we require *scale* (or size) invariance.

When applied to objects which have been segmented into *object* and *background* — $f(r,c) \in 0,1$, where 1 corresponds to *object* and 0 corresponds to *background*, moments and moment invariants capture the shape of the object silhouette.

For discrete images the two-dimensional moment of order p,q is defined as follows:

$$m_{pq} = \sum_r \sum_c r^p c^q f(r,c). \tag{1}$$

where $f(r,c)$ is the image value at row r and column c.

In the case of binary images, $f(r,c) \in 0,1$, so that $m_{00} = \sum_r \sum_c r^0 c^0 f(r,c) = \sum_r \sum_c f(r,c)$, corresponds to area measured in pixels.

Considering pixels as point masses, the centre of mass of the image (i.e. of the segmented object) is given by:

$$\bar{r} = \frac{m_{10}}{m_{00}}, \bar{c} = \frac{m_{01}}{m_{00}}. \tag{2}$$

It is easy to achieve *shift invariance* by referring all moment calculations to the centre of mass given by eqn. 2; these are the so-called *centralised moments*:

$$m'_{pq} = \sum_r \sum_c (r-\bar{r})^p (c-\bar{c})^q f(r,c). \tag{3}$$

Scale invariance is achieved by normalising with respect to area (m_{00}), i.e. if we (notionally) scale the row and column dimensions by:

$$\lambda = \frac{1}{\sqrt{m_{00}}}. \tag{4}$$

We arrive at *normalised central moments*:

$$n_{pq} = \frac{m'_{pq}}{\sqrt{m_{00}}^{\gamma}}. \tag{5}$$

where

$$\gamma = \frac{p+q}{2} + 1, p+q \geq 2. \tag{6}$$

Eqn. 6 is derived as follows. Consider a continuous image (rather than one with discrete pixels),

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r^p c^q f(r,c) dr dc. \tag{7}$$

Now scale each dimension by λ,

$$m_{\lambda,pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r^p c^q f(r/\lambda, c/\lambda) dr dc, \tag{8}$$

leading to,

$$m_{\lambda,pq} = \lambda^{2+p+q} m_{pq} \tag{9}$$

Interpretation of Moments

Here we give a brief interpretation of (*centralised*) moments, i.e. moments which are computed with respect to the centre of mass, eqn. 3:

$$m'_{pq} = \sum_r \sum_c (r-\bar{r})^p (c-\bar{c})^q f(r,c). \tag{10}$$

Let us first examine one-dimensional moments — which are often used in describing one-dimensional probability mass functions:

$$m'_x = \sum_x (x-\bar{x})^p f(x). \tag{11}$$

Here, m'_0 is the sum of $f(x)$.

m'_1 is the mean of $f(x)$ about $\bar{x}$, i.e. 0, since $m'_1 = \bar{x}$.

m'_2 is computed by weighting $f(x)$ by $(x-\bar{x})^2$, so m'_2 is a measure of the 'fatness' of $f(x)$; in probability mass functions it the *variance*.

m'_3 is computed by weighting $f(x)$ by $(x-\bar{x})^3$; note that by weighting $f(x)$ by $(x-\bar{x})^3$, pixels on the negative side of $(x-\bar{x})$ contribute negatively, whilst pixels on the positive side of $(x-\bar{x})$ contribute positively, i.e. m'_3 measures *skewness* (lack of symmetry) about $\bar{x}$. If $f(x)$ has a long tail to the right (increasing x), *positive* skewness, $m'_3 > 0$, will result; correspondingly, a long tail to the left will yield $m'_3 < 0$; and $f(x)$ symmetric about the the mean $\bar{x}$ will have $m'_3 = 0$.

Return now to two dimensions and eqn. 10. m'_{00} is the sum of $f(r,c)$, i.e. in our binary image case, it is the total area of object pixels (= 1).

m'_{10} is the r-centre of mass; $m'_{10} = 0$ since m'_{pq} are referred to the centre of mass. Likewise $m'_{01} = 0$ is is the c-centre of mass.

m'_{20} measures the 'fatness' (variance in statistics terms) in the r-dimension. Likewise m'_{02} is the fatness c-dimension. For example, long thin ellipse aligned vertically along the r-axis will have a large m'_{20} and small m'_{02}, with a long thin ellipse aligned horizontally will have large m'_{02} and small m'_{20}.

m'_{30} measures the skewness in the r-dimension and m'_{03} measures the skewness in the c-dimension. An 'egg' shape aligned vertically will have a relatively large m'_{30}; likewise the same shape aligned horizontally will have a relatively large m'_{03}. Axially symmetric shapes, for example ellipses (and circles) have $m'_{03} = m'_{30} = 0$.

m'_{11} measures the degree to which the object is aligned along the r-c diagonals; its sign indicates the proportion contributed by the four diagonals. Shapes such as ellipses aligned along one or other axis have $m'_{11} = 0$.

m'_{21} and m'_{12} provide measures of skew-symmetry, i.e. symmetry along a horizontal.

Rotational Invariance

The problem of *rotational invariance remains*. One solution, as in (Masters 1995), computes the angles of the principal axes and then recomputes $m_{\lambda,pq}$ such that moments are referred these axes. (The term *principal axes* is used to convey the same meaning as 'principal axes of an ellipse'.) Unfortunately, when principal axes are indistinct, e.g. as in a disc (circular) shape, this normalisation can sometimes fail. Hence the interest in Hu's rotationally invariant moments.

Hu's Rotationally Invariant Moments

Under scale normalisation and shift invariance, moments m_{00} (area) may be discarded, and location moments: m_{10} and m_{01} are both zero. Hu's *rotationally invariant moments* are derived from the normalised central moments of eqn. 5:

$$
\begin{aligned}
h_1 &= n_{20}+n_{02}, && (12)\\
h_2 &= (n_{20}-n_{02})^2+4n_{11}^2, && (13)\\
h_3 &= (n_{30}-3n12)^2+(3n_{21}-n_{03})^2, && (14)\\
h_4 &= (n_{30}+n_{12})^2+(n_{21}+n_{03})^2, && (15)\\
h_5 &= (n_{30}+3n12)(n_{30}+n_{12}) && (16)\\
&\quad +((n_{30}+n_{12})^2-3(n_{21}-n_{03}^2)), && (17)\\
&\quad +(3n_{21}-n_{03}) && (18)\\
&\quad (n_{21}+n_{03})(3(n_{30}+n_{12})^2-(n_{21}+n_{03})^2), && (19)\\
h_6 &= (n_{20}-n_{02})((n_{30}+n_{12})^2-(n_{21}+n_{03})^2) && (20)\\
&\quad +4n_{11}(n_{30}+n_{12})(n_{21}+n_{03}), && (21)\\
h_7 &= (3n_{21}-n_{03})((n_{30}+n_{12})(((n_{30}+n_{12})^2 && (22)\\
&\quad -3(n_{21}+n_{03})^2 && (23)\\
&\quad (3n_{21}-n_{03})(n_{21}+n_{03})(3(n_{12}+n_{30})^2-(n_{21}+n_{03})^2). && (24)
\end{aligned}
$$

The invariant moments $h_1, h_2, \ldots, h7$ are subjected to further normalisation transformations, mostly involving taking logarithms, (Masters 1995) to yield values which are of broadly comparable size and range.

References

Hu, M. (1962). Visual pattern recognition by moment invariants, *IRE Trans. Information Theory* **IT-8**: 179–187.

Masters, T. (1995). *Signal and Image Processing with Neural Networks: C++ sourcebook*, John Wiley & Sons, New York.