Fitting HyperQuadrics to range data

1 HyperQuadrics [1]

1.1 Definition

HyperQuadrics (HQ) are defined in 2D by:

$$f(x, y; \vec{\theta}) = \sum_{i=1}^{N} |a_i x + b_i y + c_i|^{\gamma_i} = 1, \tag{1}$$

where N is the any arbitrary number ($\geqslant 2$), $(x, y) \in \mathbb{R}^{+2}$ (as we are working on an image) and $\vec{\theta}$ is a $(1 \times 4N)$ vector defined by $\vec{\theta} = \{a_i, b_i, c_i, \gamma_i, i = 1..N\}$ with $(a_i, b_i, c_i, \gamma_i) \in \mathbb{R}^3 \times \mathbb{R}^{*+}$.

Each term in the summation (1) is positive and the terms add to one. Therefore, no individual term can ever exceed one. Since the γ_is are positive, this implies that $a_i x + b_i y + c_i \leqslant 1$ $\forall i = 1, 2, \cdots, N$. Thus each term gives a pair of parallel bounding lines. The parameters a_i and b_i in term i give the slope of the lines corresponding to that term. The distance S_i between the bounding lines is given by $2/\sqrt{a_i^2 + b_i^2}$.

The inside-outside function of the hyperquadric F_{io} is then defined by:

$$F_{io}(x, y; \vec{\theta}) = \left(\sum_{i=1}^{N} |a_i x + b_i y + c_i|^{\gamma_i} \right)^p, \tag{2}$$

where p is usually taken equal to 1/5, since the γ_is are usually around 5. The inside-outside function gives a measure of the relative position of a data point with respect to the model and it can be used as an error-of-fit (EoF) measure in the shape recovery procedure:

$$EoF(\vec{\theta}) = \sum_{i=1}^{N_{data}} \frac{1}{\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2} (1 - F_{io}(x_i, y_i; \vec{\theta}))^2. \tag{3}$$

1.2 Adding constraints

The EoF function has an infinite number of global minima with zero error and only some of them correspond to the shape being modeled. This is in addition to the existence of many local minima that might produce shapes very different from the desired shape. Clearly, some restrictions on the range of values the parameters can take are needed. Therefore the lower limit on the distance S_i between a pair of bounding lines is set to be slightly larger than the minimum diameter S_{min} of the object. The upper limit on S_i is set slightly larger the maximum diameter S_{max} of the object. Thus we have $\forall i = 1, 2, \cdots N$,

$$k_2 S_{min} \leqslant S_i \leqslant k_1 S_{max} \tag{4}$$

$$\Leftrightarrow \quad \mu_1 \leqslant a_i^2 + b_i^2 \leqslant \mu_2, \tag{5}$$

where $\mu_1 = (1/(k_1 S_{max}))^2$ and $\mu_2 = (1/(k_2 S_{min}))^2$ with k_1 and k_2 some constants.

Next, some restrictions on the γ_is are added. When γ_i is very high, it has a tendency to increase more and more. Also when the γ_is are smaller than 1, the shape has concavities. Therefore the γ_is are restricted to lie in the interval [1; 15]. This gives two more constraints: $1 \leqslant \gamma_i \leqslant 15$.

Given a single inequality constraint $g(x, y) \leqslant 0$, a penalty function can be defined as $P(x, y) = (max[0, g(x, y)])^2$. The penalty function P_i for each term becomes:

$$P_i(\vec{\theta}) = (max[0, \mu_1 - (a_i^2 + b_i^2)])^2 + (max[0, a_i^2 + b_i^2 - \mu_2])^2 + (max[0, 1 - \gamma_i])^2 + (max[0, \gamma_i - 30])^2. \quad (6)$$

And the EoF function becomes

$$EoF(\vec{\theta}) = \sum_{i=1}^{N_{data}} \frac{1}{\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2}(1 - F_{io}(x_i, y_i; \vec{\theta})^2 + \nu \sum_{i=1}^{N} P_i(\vec{\theta}), \quad (7)$$

where ν is a positive constant.

1.3 Minimization of the EoF function

An iterative algorithm is used to minimized the EoF (7):

1. minimize EoF and obtain F_{io};
2. compute the weight for each data point: $w_i = 1/\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2$;
3. use the weight w_i computed in 2) to minimize EoF, and obtain F_{io};
4. if the fitting F_{io} is not good enough, go to 2); else, stop.

In the above procedure, the minimization is accomplished by the Levenberg-Marquardt non-linear optimization method. This method need the first order derivative of the EoF with respect to the fit parameters.

2 Levenberg-Marquardt method

The presentation of the Levenberg-Marquardt (LM) non-linear optimization method will follow the one proposed in [2] applied to HQ.

2.1 Principle

The model to be fitted is (2), and the χ^2 merit function is:

$$\chi^2(\vec{\theta}) = \sum_{i=1}^{N_{data}} \frac{1}{\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2}(1 - F_{io}(x_i, y_i; \vec{\theta}))^2 + \nu \sum_{j=1}^{N} P_j(\vec{\theta}). \quad (8)$$

The gradient of χ^2 with respect to the parameters $\vec{\theta}$, which will be zero at the χ^2 minimum, has components

$$\frac{\partial \chi^2}{\partial \theta_k}(\vec{\theta}) = -2 \sum_{i=1}^{N_{data}} \frac{1}{\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2}(1 - F_{io}(x_i, y_i; \vec{\theta}))\frac{\partial F_{io}(x_i, y_i; \vec{\theta})}{\partial \theta_k} + \nu \sum_{j=1}^{N} \frac{\partial P_j}{\partial \theta_k}(\vec{\theta}), \quad (9)$$

where the term $\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2$ is considered constant (as in [1]).

Then taking an additional derivative and neglecting the second order derivative (see p.683 in [2] for explaination) gives

$$\frac{\partial^2 \chi^2}{\partial\theta_k \partial\theta_l}(\vec{\theta}) = 2\sum_{i=1}^{N_{data}} \frac{1}{\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2} \frac{\partial F_{io}(x_i, y_i; \vec{\theta})}{\partial\theta_k} \frac{\partial F_{io}(x_i, y_i; \vec{\theta})}{\partial\theta_l} + \nu \sum_{j=1}^{N} \frac{\partial^2 P_j}{\partial\theta_k \partial\theta_l}(\vec{\theta}). \tag{10}$$

See section 2.2 and 2.3 for the derivation of the different terms.

Defining

$$\beta_k \equiv -\frac{1}{2}\frac{\partial \chi^2}{\partial\theta_k} \equiv \sum_{i=1}^{N_{data}} \frac{1}{\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2}(1 - F_{io}(x_i, y_i; \vec{\theta}))\frac{\partial F_{io}(x_i, y_i; \vec{\theta})}{\partial\theta_k} - \frac{1}{2}\nu \sum_{j=1}^{N} \frac{\partial P_j}{\partial\theta_k}(\vec{\theta}) \tag{11}$$

$$\alpha_{kl} \equiv \frac{1}{2}\frac{\partial^2 \chi^2}{\partial\theta_k \partial\theta_l} \equiv \sum_{i=1}^{N_{data}} \frac{1}{\|\nabla F_{io}(x_i, y_i; \vec{\theta})\|^2} \frac{\partial F_{io}(x_i, y_i; \vec{\theta})}{\partial\theta_k} \frac{\partial F_{io}(x_i, y_i; \vec{\theta})}{\partial\theta_l} + \frac{1}{2}\nu \sum_{j=1}^{N} \frac{\partial^2 P_j}{\partial\theta_k \partial\theta_l}(\vec{\theta}) \tag{12}$$

and

$$\alpha'_{kl} \equiv \begin{cases} \alpha_{kl}(1+\lambda), & \text{if } k = l, \\ \alpha_{kl}, & \text{if } k \neq l, \end{cases} \tag{13}$$

the variation of the parameters $\delta\alpha_k$ are computed according to

$$\sum_{l=1}^{4N} \alpha'_{kl}\delta\theta_l = \beta_k \qquad \text{or} \qquad [\alpha']\delta\vec{\theta} = \vec{\beta}. \tag{14}$$

Given an initial guess for the set of fitted parameters $\vec{\theta}$, the minimization process is as follows:

- Compute $\chi^2(\vec{\theta})$;
- Pick a modest value for λ, say $\lambda = 0.05$;
- (†) Solve the linear equations (14) for $\delta\vec{\theta}$ and evaluate $\chi^2(\vec{\theta} + \delta\vec{\theta})$;
- if $\chi^2(\vec{\theta} + \delta\vec{\theta}) \geqslant \chi^2(\vec{\theta})$, increase λ by a factor of 50 (or any other substantial factor) and go back to (†);
- if $\chi^2(\vec{\theta} + \delta\vec{\theta}) < \chi^2(\vec{\theta})$, decrease λ by a factor of 50, update the trial solution $\vec{\theta} \leftarrow \vec{\theta} + \delta\vec{\theta}$, and go back to (†).

2.2 Derivatives of the HQ function

First consider the HQ equation (1). Its first order derivative with respect to the HQ parameters are:

$$\frac{\partial f}{\partial a_i}(x, y; \vec{\theta}) = \gamma_i x \cdot \operatorname{sign}(a_i x + b_i y + c_i) \cdot |a_i x + b_i y + c_i|^{\gamma_i - 1}, \tag{15}$$

$$\frac{\partial f}{\partial b_i}(x, y; \vec{\theta}) = \gamma_i y \cdot \operatorname{sign}(a_i x + b_i y + c_i) \cdot |a_i x + b_i y + c_i|^{\gamma_i - 1}, \tag{16}$$

$$\frac{\partial f}{\partial c_i}(x, y; \vec{\theta}) = \gamma_i \cdot \operatorname{sign}(a_i x + b_i y + c_i) \cdot |a_i x + b_i y + c_i|^{\gamma_i - 1}, \tag{17}$$

as $\gamma_i > 0$.

$$\frac{\partial f}{\partial \gamma_i}(x, y; \vec{\theta}) = \frac{\partial}{\partial \gamma_i}\exp(\ln(|a_i x + b_i y + c_i|^{\gamma_i})) = \frac{\partial}{\partial \gamma_i}\exp(\gamma_i \ln(|a_i x + b_i y + c_i|))$$

$$\Leftrightarrow \quad \frac{\partial f(x, y; \vec{\theta})}{\partial \gamma_i} = \ln(|a_i x + b_i y + c_i|) \cdot |a_i x + b_i y + c_i|^{\gamma_i}. \tag{18}$$

And the derivative of (2) wrt. the HQ parameters is:

$$\frac{\partial F_{io}(x, y; \vec{\theta})}{\partial\theta_k} = \frac{1}{5}\frac{\partial f(x, y; \vec{\theta})}{\partial\theta_k} f(x, y; \vec{\theta})^{-\frac{4}{5}}. \tag{19}$$

2.3 Derivatives of the penalty term

Now considering the penalty term of the EoF function (6), its first order derivative wrt. the HQ parameters are given by:

$$\frac{\partial P_i}{\partial a_i}(\vec{\theta}) = 4a_i(-max[0, \mu_1 - (a_i^2 + b_i^2)] + max[0, a_i^2 + b_i^2 - \mu_2]), \tag{20}$$

$$\frac{\partial P_i}{\partial b_i}(\vec{\theta}) = 4b_i(-max[0, \mu_1 - (a_i^2 + b_i^2)] + max[0, a_i^2 + b_i^2 - \mu_2]), \tag{21}$$

$$\frac{\partial P_i}{\partial \gamma_i}(\vec{\theta}) = -2max[0, 1 - \gamma_i] + 2max[0, \gamma_i - 15]. \tag{22}$$

$$\frac{\partial P_i}{\partial \theta_k}(\vec{\theta}) = 0 \qquad \text{otherwise.} \tag{23}$$

The second derivative and cross derivative are given by:

$$\begin{aligned}\frac{\partial^2 P_i}{\partial a_i^2}(\vec{\theta}) = 4(&-max[0, \mu_1 - (a_i^2 + b_i^2)] + max[0, a_i^2 + b_i^2 - \mu_2] \\ &+ 2a_i^2(\text{sign}(max[0, \mu_1 - (a_i^2 + b_i^2)]) + \text{sign}(max[0, a_i^2 + b_i^2 - \mu_2]))),\end{aligned} \tag{24}$$

$$\begin{aligned}\frac{\partial^2 P_i}{\partial b_i^2}(\vec{\theta}) = 4(&-max[0, \mu_1 - (a_i^2 + b_i^2)] + max[0, a_i^2 + b_i^2 - \mu_2] \\ &+ 2b_i^2(\text{sign}(max[0, \mu_1 - (a_i^2 + b_i^2)]) + \text{sign}(max[0, a_i^2 + b_i^2 - \mu_2]))),\end{aligned} \tag{25}$$

$$\frac{\partial^2 P_i}{\partial a_i \partial b_i}(\vec{\theta}) = 8a_i b_i(\text{sign}(max[0, \mu_1 - (a_i^2 + b_i^2)]) + \text{sign}(max[0, a_i^2 + b_i^2 - \mu_2])) \tag{26}$$

$$\frac{\partial^2 P_i}{\partial \gamma_i^2}(\vec{\theta}) = 2(\text{sign}(max[0, 1 - \gamma_i]) + \text{sign}(max[0, \gamma_i - 15])), \tag{27}$$

$$\frac{\partial^2 P_i}{\partial \theta_k \partial \theta_l}(\vec{\theta}) = 0 \qquad \text{otherwise.} \tag{28}$$

References

[1] S. Kumar and S. Han and D. Goldgof and K. Bowyer, *On recovering hyperquadrics from range data, IEEE Trans. Pattern Anal. Machine Intell.*, Vol 17, pp. 1079 – 1083, 1995.

[2] W.H. Press and S.A. Teukolsky and W.T. Vetterling B.P. Flannery, *Numerical Recipes in C*, Cambridge, England, Cambridge Univ. Press, 1992.