Dual Approaches for Elliptic Hough Transform: Eccentricity/Orientation vs Center based

Philippe Latour and Marc Van Droogenbroeck

University of LIEGE, Montefiore Research Institute

Post-print version with error corrections (March 29, 2019)

Abstract

Ellipse matching is the process of extracting (detecting and fitting) elliptic shapes from digital images. It appears that it requires the determination of typically five parameters, which can be determined by using an Elliptic Hough Transform (EHT) algorithms.

In this paper, we focus on Elliptic Hough Transform (EHT) algorithms based on two edge points and their associated image gradients. For this scenario, authors first reduce the dimension of the 5D EHT by means of some geometrical observations, and then apply a simpler HT. We present an alternative approach, more specifically an algebraic framework, based on the pencil of bi-tangent conics, expressed in two dual forms: the point or the tangential forms. It appears that, for both forms, the locus of the ellipse parameters is a line in a 5D space.

With this framework, we can split the EHT into two steps. The first step accumulates 2D lines, which are computed from planar projections of the parameter locus (5D line). The second part back-projects the peak of the 2D accumulator into the 5D space, to obtain the three remaining parameters that we then accumulate in a 3D histogram, possibly represented as three separated 1D-histograms.

For the point equation, the first step extracts parameters related to the ellipse orientation and eccentricity, while the remaining parameters are related to the center and a sizing parameter of the ellipse. For the tangential equation, the first process is the known center extraction algorithm, while the remaining parameters are related to the ellipse half-axes and orientation.

keywords: ellipse detection, ellipse matching, Hough Transform, pencil of conics, tangential equation.

1 Introduction

Ellipses are planar closed curves described by five independent geometric parameters (or six algebraic coefficients, defined up to a scaling factor). They are ubiquitous in images, especially those representing human environments, and they help in describing the scene. Ellipse extraction in images is a particular case of template matching, which is generally solved either by area-based or by feature-based approaches. This paper investigates specific feature-based approaches, where the chosen features are the edge points (or contour points) extracted from the image.

Indeed, if most of an ellipse is visible in the image, we expect edge points to be localizable on the ellipse border and the image gradient at these edge points to be directed along the normal to the ellipse contour (and then also orthogonal to the tangent to the ellipse). So, edge points with their associated gradients are possible evidences of the presence of an ellipse in the image.

Conceptually, template (ellipse) matching relies on two interdependent parts; detection (selection of the template in the image) and fitting (estimation of the template parameters). Ellipse fitting algorithms (see [1]) estimate well the parameters of an unknown ellipse from a set of points originating from this ellipse, but they are absolutely not robust to outliers. So, in practical applications, we first need to select or detect in images the points belonging to an ellipse before being able to fit it.

There exist two main paradigms for ellipse detection; RANSAC-like methods (see [1], chapter 4) and Hough Transform (HT) algorithms (see [2] and [3]). Both approaches have assets and drawbacks, and may even be intermixed in practical applications, but this paper will only consider the HT. Indeed, we focuses on a specific contour-based *ellipse matching* method named *Elliptic Hough Transform* (EHT), where the contour is represented by a list of edge points extracted from the image.

In HT, the unknown template (ellipse) is represented by a point in its parameter space (five-dimensional space for ellipses) and all the evidences of the presence of the template in the image (edge points) highlights/votes for some subset of this parameter space. For instance, an edge point and its associated gradient impose two constraints to the unknown ellipse (passing by the point and touching the line orthogonal to the gradient at this point). So, these two evidences reduce the compatible ellipse parameter set to a three-dimensional space (five-dimensional parameter space minus two constraints). We say that these two evidences votes for the compatible three-dimensional parameter subset. Then, in the parameter space, we accumulate the votes of all the evidences (edge points and gradients) extracted from the image. At the end of the process, the algorithm chooses (extracts) the template (ellipse) that has accumulated the largest number of votes.

EHT methods aim, first, at detecting the presence of any number of elliptic structures in an image and, incidentally, at computing an approximation of their geometric parameters or algebraic coefficients. Obviously, when the edge points belonging to an ellipse has been extracted by the EHT, if the approximated geometric parameters are not accurate enough, we are free to apply any additional ellipse fitting algorithm.

In the literature, authors use different types and number of evidences for EHT. In this paper, the elementary evidence groups are *pairs of edge points with their associated gradients,* and we impose no conditions on the gradients. In the following, this evidence group is named "a pair of edge points" for simplicity; the availability of associated gradients is implicitly assumed. A pair of edge point imposes four constraints on the compatible ellipses (passing by two points and touching two lines). Therefore, a pair of edge points votes for a one-dimensional ellipse parameter subset that is compatible with this evidence group.

In EHT, we would need to accumulate the votes in a five-dimensional space, which is complicated in practice. So, authors separate the problem in two steps. In [4], Tsuji et al. presented an EHT algorithm based on pairs of edge points with parallel gradients, which use a kind of two-dimensional histogram for computing the center of the ellipse. But, Yuen et al. [5] were the first to use pairs of edge points with arbitrary gradients. They accumulates lines in the 2D Hough space of the ellipse centers and then accumulate planes in the 3D Hough space of the remaining normalized algebraic coefficients. Other authors (see [6] and [7]) use the same first step but resort to specific geometric parameters for the three remaining parameters, which they determine in two steps rather than one.

It is interesting to note that two articles, more specifically [8] and [9], presented, independently and without explicitly naming it, the pencil of bi-tangent conics as the right mathematical framework for analyzing the family of ellipses built on two edge points. In fact, Yoo et al. [8] use the framework only for the first part of the EHT algorithm. Besides, Benett et al. [9] noticed that the framework enables to compute, as a first step, any pair or triplet of parameters (not only the center) and to easily incorporate any prior geometrical constraints. None of these works consider the framework for the second step of the EHT.

A key point of the EHT is the choice of the set of (algebraic or geometric) parameters used for the accumulation process. In the following, we show that the usual method of center accumulation is well described within the framework of the the dual equation. We also show that the set of parameters of the direct equation of the framework was introduced in [10] for fitting applications (see [11], [12]), but has never been used explicitly in the accumulation process of the EHT. This direct form accumulates a pair of parameters, which are equivalent to the eccentricity and orientation of the ellipse.

2 Problem Statement

In this section, we introduce the main notation and equations. In the following, a point, named P, will be represented by the $2D$ vector of its Cartesian coordinates $\bar{\boldsymbol{P}} = \begin{pmatrix} X_P & Y_P \end{pmatrix}^T$ and/or by the $3D$ vector of its homogeneous coordinates $\boldsymbol{P} = \begin{pmatrix} x_P & y_P & z_P \end{pmatrix}^T$ or $\boldsymbol{P} = \begin{pmatrix} X_P & Y_P & 1 \end{pmatrix}^T$. Likewise, an unknown point will be represented by its Cartesian coordinate $\bar{\boldsymbol{X}} = \begin{pmatrix} X & Y \end{pmatrix}^T$ and also by its homogeneous coordinates $\boldsymbol{X} = \begin{pmatrix} x & y & z \end{pmatrix}^T$ or $\boldsymbol{X} = \begin{pmatrix} X & Y & 1 \end{pmatrix}^T$. A line, named l, will then satisfy an equation $\boldsymbol{l}^T\boldsymbol{X} = 0$, where $\boldsymbol{l} = \begin{pmatrix} u_l & v_l & w_l \end{pmatrix}^T$ is the $3D$ vector of the homogeneous coefficients of its equation. Also, an unknown line will satisfy an equation $\boldsymbol{u}^T\boldsymbol{X} = 0$, where $\boldsymbol{u} = \begin{pmatrix} u & v & w \end{pmatrix}^T$. Most of the time, we will use the homogeneous coordinates of the projective plane $\mathbb{P}^2 = P^2(\mathbb{R})$ and, by an abuse of language, we will often identify the geometric elements with their representation.

2.1 Ellipses and Conics

Geometric parameters of ellipses

Geometrically speaking, an ellipse is usually defined by (the Cartesian coordinates of) its center (C, C'), its half-axes (r, r'), and its orientation θ. In the following, we will also use the eccentricity e, defined as $e^2 = 1 - {r'^2}/{r^2}$.

Point (direct) equation

An ellipse may be defined algebraically as a special kind of conics, which is a second degree algebraic curve. A conic $\mathcal{C}$ is any curve whose point coordinates satisfy the following equation

$$\mathcal{C} : aX^2 + 2b''XY + a'Y^2 + 2b'X + 2bY + a'' = \boldsymbol{X}^T\mathtt{F}\boldsymbol{X} = 0\,, \tag{1}$$

where the matrix $\mathtt{F}$ is easily obtained by identification of the two members in equation (1).

The conic is also fully determined by the vector of its six homogeneous coefficients $\boldsymbol{f} = (a, b'', a', b', b, a'')^T \in \mathbb{R}^6 \setminus \{\boldsymbol{0}\}$, defined up to a scaling factor. The conic may then be represented by a point in the projective space $\mathbb{P}^5 = P^5(\mathbb{R})$. In the following, we will use the operator *flatten* that operates on the symmetric matrix $\mathtt{F}$ and returns the vector $\boldsymbol{f}$, so we write $\boldsymbol{f} = \text{flatten}(\mathtt{F})$.

We will later show (see equation (8)) that, when the conic is an ellipse, we always have $a + a' \neq 0$ and we may also use the following alternative form (introduced in [10]) for the algebraic equation of an ellipse:

$$\left(X^2 + Y^2\right) - d\left(X^2 - Y^2\right) - 2d'XY - 2cX - 2c'Y - c'' = 0\,, \tag{2}$$

where the five *direct* coefficients $\bar{\boldsymbol{f}} = (d, d', c, c', c'')$ are related to the homogeneous direct coefficients in equation (1) and to the geometric parameters by

$$\left\{\begin{array}{lllll} d & = & \frac{a'-a}{a+a'} & = & \frac{e^2\cos 2\theta}{2-e^2} \\ d' & = & \frac{-2b''}{a+a'} & = & \frac{e^2\sin 2\theta}{2-e^2} \end{array}\right. ,\quad \left\{\begin{array}{lllll} c & = & \frac{-2b'}{a+a'} & = & (1-d)C - d'C' \\ c' & = & \frac{-2b}{a+a'} & = & -d'C + (1+d)C' \\ c'' & = & \frac{-2a''}{a+a'} & = & \frac{2r^2r'^2}{r^2+r'^2} - cC - c'C'\,. \end{array}\right. \tag{3}$$

The notation $\bar{\boldsymbol{f}}$ for the $5D$ vector and $\boldsymbol{f}$ for the $6D$ homogeneous vector is similar to the notation $\bar{\boldsymbol{P}}$ for the $2D$ Cartesian coordinates and $\boldsymbol{P}$ for the $3D$ homogeneous coordinates. In some way, the vectors $\boldsymbol{f}$ and $\bar{\boldsymbol{f}}$ may be respectively considered as the projective and Cartesian direct coordinates of an ellipse.

Tangential (dual) equation

A line $\boldsymbol{u}^T\boldsymbol{X} = 0$ is a tangent to the conic described by the equation (1) if and only if its triplet of coefficients $\boldsymbol{u}$ satisfies the dual equation (also called *tangential equation*) of the conic

$$\mathcal{C}^* : Au^2 + 2B''uv + A'v^2 + 2B'uw + 2Bvw + A''w^2 = \boldsymbol{u}^T\mathtt{F}^*\boldsymbol{u} = 0\,, \tag{4}$$

where $\mathtt{F}^*$ is the matrix of the cofactors of $\mathtt{F}$. And, we also represent the dual conic by the vector of its six homogeneous coefficients $\boldsymbol{F} = (A, B'', A', B', B, A'')^T = \text{flatten}(\mathtt{F}^*)$.

For an ellipse, $A'' = aa' - b''^2 \neq 0$ and we also have an alternative form for the dual equation:

$$D''\left(u^2 + v^2\right) + 2D'uv - D\left(u^2 - v^2\right) + 2Cuw + 2C'vw + w^2 = 0\,, \tag{5}$$

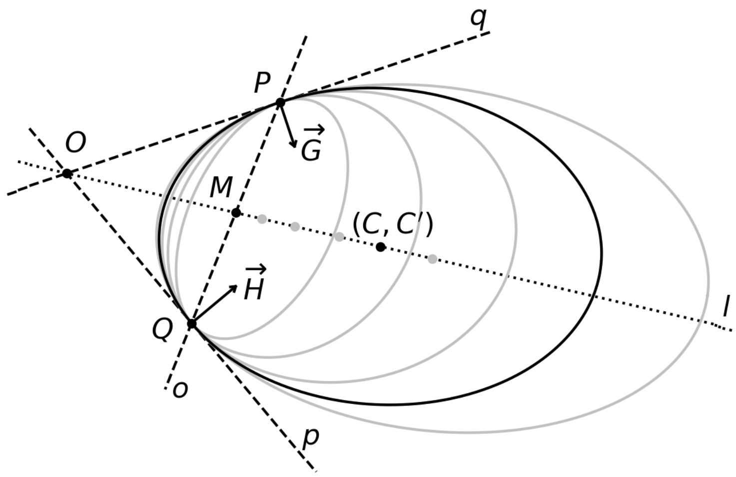

Figure 1: The edge points P, Q, and their associated gradients $\overrightarrow{\mathrm{G}}$, $\overrightarrow{\mathrm{H}}$ are evidences of the presence of an ellipse in the image.

where the 5 *dual* coefficients $\bar{\boldsymbol{F}} = (C, C', D, D', D'')$ are related to the dual homogeneous coefficients in the equation (4) and to the geometric parameters by

$$\left\{\begin{array}{lcl} C & = & \frac{B'}{A''} \\ C' & = & \frac{B}{A''} \end{array}\right. \quad , \quad \left\{\begin{array}{lclcl} D & = & \frac{A'-A}{2A''} & = & \frac{1}{2}\left(C'^2-C^2\right)+\frac{1}{2}\left(r^2-r'^2\right)\cos 2\theta \\ D' & = & \frac{B''}{A''} & = & CC'-\frac{1}{2}\left(r^2-r'^2\right)\sin 2\theta \\ D'' & = & \frac{A+A'}{2A''} & = & \frac{1}{2}\left(C^2+C'^2-r^2-r'^2\right). \end{array}\right. \tag{6}$$

We highlight the fact that the two coefficients (C, C') in the equation (6) effectively represents the Cartesian coordinates of the ellipse center and have thus the same notation. The vectors $\boldsymbol{F}$ and $\bar{\boldsymbol{F}}$ may be respectively considered as the projective and Cartesian dual coordinates of an ellipse.

Type of a conic

A conic is an ellipse if and only if

$$A'' = aa' - b''^2 > 0 \quad \text{or} \quad 0 \leq e < 1\,. \tag{7}$$

This condition implies that $aa' > b''^2 \geq 0$. Subsequently, a and a' must have the same sign and may not be null. Then, we also have the following relations

$$\operatorname{sign}(a) = \operatorname{sign}(a') = \operatorname{sign}(a + a') \quad \text{and} \quad a + a' \neq 0\,. \tag{8}$$

With the notation $\Delta = \det \mathtt{F}$, an ellipse is a real ellipse if, in addition,

$$a\Delta < 0 \Longleftrightarrow (a + a')\,\Delta < 0\,. \tag{9}$$

2.2 Elliptic Features - Triangle of Interest

We consider two edge points P and Q and their associated gradients $\overrightarrow{\mathrm{G}}$ and $\overrightarrow{\mathrm{H}}$; this configuration is illustrated in Figure 1.

The mid-point of P and Q is represented by $\boldsymbol{M} = \frac{1}{2}(\boldsymbol{P} + \boldsymbol{Q})$, where $\boldsymbol{P} = \begin{pmatrix} X_P & Y_P & 1 \end{pmatrix}^T$, $\boldsymbol{Q} = \begin{pmatrix} X_Q & Y_Q & 1 \end{pmatrix}^T$ and we define $\boldsymbol{K} = \boldsymbol{Q} - \boldsymbol{P}$. The line o (with $\boldsymbol{o} = \boldsymbol{P} \wedge \boldsymbol{Q}$), joining the points $\boldsymbol{P}$ and $\boldsymbol{Q}$, is a chord secant to the unknown ellipse.

The vectors $\overrightarrow{\mathrm{G}}$ and $\overrightarrow{\mathrm{H}}$ are the image gradients at the points P and Q. Their direction is represented by the homogeneous vectors $\boldsymbol{G} = \begin{pmatrix} X_G & Y_G & 0 \end{pmatrix}^T$ and $\boldsymbol{H} = \begin{pmatrix} X_H & Y_H & 0 \end{pmatrix}^T$. We also introduce the dot products $\lambda_G = \boldsymbol{G} \cdot \boldsymbol{K}$ and $\lambda_H = \boldsymbol{H} \cdot \boldsymbol{K}$ and we will later show that they cannot be null in the practical situations of interest.

The line q, defined by $\boldsymbol{q} = \frac{1}{\lambda_G}\boldsymbol{P} \wedge (\boldsymbol{G} \wedge \boldsymbol{e_z})$ with $\boldsymbol{e_z} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$, passing by P and orthogonal to $\overrightarrow{\mathrm{G}}$, is tangent to the unknown ellipse. Likewise, the line p (with $\boldsymbol{p} = \frac{1}{\lambda_H}(\boldsymbol{H} \wedge \boldsymbol{e_z}) \wedge \boldsymbol{Q}$) passing by Q and orthogonal to $\overrightarrow{\mathrm{H}}$, is tangent to the unknown ellipse. In the expression of $\boldsymbol{p}$ and $\boldsymbol{q}$, we choose the arbitrary scaling factors λ_G and λ_H to simplify the expression of some of the following equations. The intersection point O of the two lines q and p is represented by $\boldsymbol{O} = \boldsymbol{p} \wedge \boldsymbol{q}$.

We consider the triangle $\triangle$OPQ defined by the edge points P, Q and the intersection point O. The sides of this triangle are thus the lines o, q and p. The matrix of the triangle vertices is named $\mathtt{V}$. Thanks to the choice of the scaling factors λ_G and λ_H, it can be shown that the matrix $\mathtt{S} = \mathtt{V}^*$ of the cofactors of $\mathtt{V}$ is also the matrix of the coefficients of the side lines of the triangle. Therefore, we have

$$\mathtt{V} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{P} & \boldsymbol{Q} \end{pmatrix} \quad \text{and} \quad \mathtt{V}^* = \mathtt{S} = \begin{pmatrix} \boldsymbol{o} & \boldsymbol{p} & \boldsymbol{q} \end{pmatrix}, \tag{10}$$

with $\det \mathtt{S} = \det \mathtt{V} = 1$, $\mathtt{V}^{-1} = \mathtt{S}^T$ and $\mathtt{S}^{-1} = \mathtt{V}^T$.

Conditions on the edge points

In Figure 1, we see that it is only possible to build an ellipse from a pair of edge points P and Q if they are distinct, if their gradients $\overrightarrow{\mathrm{G}}$ and $\overrightarrow{\mathrm{H}}$ are not null and if the edge point Q (resp. P) do not belongs to q (resp. p). So, in practice, we have to discard the pairs of edge points that would not verify the corresponding conditions

$$\begin{cases} \boldsymbol{K} \neq \boldsymbol{0}, \boldsymbol{G} \neq \boldsymbol{0}, \boldsymbol{H} \neq \boldsymbol{0} \\ \lambda_G = \boldsymbol{G} \cdot \boldsymbol{K} \neq 0, \lambda_H = \boldsymbol{H} \cdot \boldsymbol{K} \neq 0. \end{cases} \tag{11}$$

2.3 Pencil of Bi-Tangent Conics

The unknown ellipse belongs to the family of conics passing through the two points P and Q and tangent to the lines q and p. Mathematically speaking, this family is the pencil of conic bi-tangent to the lines q and p at their respective intersection points (P and Q) with the line o.

The general equation of a conic $\mathcal{C}_k$ belonging to this pencil is

$$\mathcal{C}_k : \left(\boldsymbol{p}^T\boldsymbol{X}\right)\left(\boldsymbol{q}^T\boldsymbol{X}\right) - \frac{k}{2}\left(\boldsymbol{o}^T\boldsymbol{X}\right)^2 = 0, \tag{12}$$

where k is any real number, which is named the *index* of the conic $\mathcal{C}_k$ in the pencil.

As said before, a conic is fully described by five parameters. As all the conics in the pencil have to verify four constraints (to pass through two given points and to touch two given lines), we obtain an equation for a conic in the pencil that only depends on one parameter, namely the index k. The notion of index of the conic is an essential component of our approach in Section 3.

Point (direct) equation of the pencil

From equation (12), we compute the matrix of a conic in the pencil $\mathtt{F_k} = \boldsymbol{p}\boldsymbol{q}^T + \boldsymbol{q}\boldsymbol{p}^T - k\boldsymbol{o}\boldsymbol{o}^T$ and also $\mathtt{F_k} = \mathtt{S}\begin{pmatrix} -k\boldsymbol{e_x} & \boldsymbol{e_z} & \boldsymbol{e_y} \end{pmatrix}\mathtt{S}^T$, where $\mathtt{S}$ is the matrix defined in equation (10). By identification with equations (1), we obtain the corresponding vector of direct homogeneous coefficients $\boldsymbol{f_k} = \boldsymbol{f_{pq}} - k\boldsymbol{f_{oo}}$, where $\boldsymbol{f_{pq}} = \text{flatten}\left(\boldsymbol{p}\boldsymbol{q}^T + \boldsymbol{q}\boldsymbol{p}^T\right)$ and $\boldsymbol{f_{oo}} = \text{flatten}\left(\boldsymbol{o}\boldsymbol{o}^T\right)$.

From (3), if we pose $\alpha_{pq} = u_p u_q + v_p v_q$ and $\alpha_{oo} = u_o^2 + v_o^2$, we obtain the direct Cartesian coefficients of the ellipses in the pencil, where $\bar{\boldsymbol{f}}_{\boldsymbol{pq}}$ and $\bar{\boldsymbol{f}}_{\boldsymbol{oo}}$ may easily be computed from (3),

$$\bar{\boldsymbol{f}}_{\boldsymbol{k}} = \xi_k \bar{\boldsymbol{f}}_{\boldsymbol{pq}} - (1 + \alpha_{pq}\xi_k)\,\bar{\boldsymbol{f}}_{\boldsymbol{oo}}, \text{ with } \xi_k = \frac{2}{-2\alpha_{pq} + k\alpha_{oo}}. \tag{13}$$

Tangential (dual) equation of a conic in the pencil

Likewise, the dual or tangential equation is determined by the matrix of the cofactors of the matrix of the conic and may be expressed (when the conic is not degenerated) in two ways $\mathtt{F}_\mathtt{k}^* =$

$\mathtt{V}\begin{pmatrix} -\boldsymbol{e_x} & k\boldsymbol{e_z} & k\boldsymbol{e_y} \end{pmatrix}\mathtt{V}^T$ and $\mathtt{F}^*_\mathtt{k} = -\boldsymbol{OO}^T + k\left[\boldsymbol{PQ}^T + \boldsymbol{QP}^T\right]$. By identification with equations (4), we obtain the corresponding vector of dual homogeneous coefficients $\boldsymbol{F_k} = -\boldsymbol{F_{OO}} + k\boldsymbol{F_{PQ}}$, where $\boldsymbol{F_{PQ}} = \text{flatten}\left(\boldsymbol{PQ}^T + \boldsymbol{QP}^T\right)$ and $\boldsymbol{F_{OO}} = \text{flatten}\left(\boldsymbol{OO}^T\right)$.

And, from (6), we obtain the dual Cartesian coefficients of the ellipses in the pencil, where $\bar{\boldsymbol{F}}_{\boldsymbol{PQ}}$ and $\bar{\boldsymbol{F}}_{\boldsymbol{OO}}$ may easily be computed from (6),

$$\bar{\boldsymbol{F}}_{\boldsymbol{k}} = \left(1 + \Xi_k {z_O}^2\right)\bar{\boldsymbol{F}}_{\boldsymbol{PQ}} - \Xi_k \bar{\boldsymbol{F}}_{\boldsymbol{OO}}\,, \text{ with } \Xi_k = \frac{1}{2k - z_O^2}\,. \tag{14}$$

Type of a conic in the pencil

Because we are looking for a real ellipse, from equations (7) and (9), we should have $2k > {z_O}^2 \Leftrightarrow \Xi_k > 0$ and $k > 2\frac{\alpha_{pq}}{\alpha_{oo}} \Leftrightarrow \xi_k > 0$ and it is possible to demonstrate that the second condition is always verified when the first one is.

3 An Algorithm for Obtaining the Matching Ellipse

3.1 Parameters of an Ellipse built upon Two Edge Points and their Gradients

We may identify the set of all the conics in the plane with the projective space $\mathbb{P}^5 = P^5(\mathbb{R})$ of their algebraic coefficients. The subset of the (algebraic coefficients of the) conics compatible with one evidence group (a pair of edge points) is a pencil of bi-tangent conics. From equation (12), we derive that such a pencil of conics is represented by a line in $\mathbb{P}^5$.

Unfortunately, the scale and the domain of the algebraic coefficients are not easy to determine because these coefficients are only defined up to a scaling factor. So, it would be more practical to accumulate in the space of the geometric parameters. But, in such a space, the pencil of conics is represented by a much more complicated curve.

A solution consists to use equation (13) in the $5D$ space of the parameters (d, d', c, c', c'') or equation (14) in the dual $5D$ space of the parameters (C, C', D, D', D''). In both spaces, the pencil of bi-tangent conics is again a line and the domain of variation of the parameters is easy to determine for a practical application. Indeed, these parameters may be easily expressed in terms of the geometric parameters of the ellipse.

3.2 A General Algorithm

The general principles and steps of a typical algorithm for the computation of the Elliptic Hough Transform (EHT) are:

- We define a $5D$ discrete accumulator corresponding to the $5D$ space of the parameters (d, d', c, c', c'') or (C, C', D, D', D''). In this accumulator, each cell is indexed by a $5D$ vector whose components are the quantized values of the corresponding parameters. All cells are initialized to 0.
- For all pairs of edge points which verify the conditions (11), in the $5D$ accumulator, we draw the line corresponding to the pencil of bi-tangent conics.
- When all the evidences has been processed and the corresponding lines drawn, we look for the peak (maximum) values in the accumulator.
- If the value of this peak, which represents the number of votes for a vector of parameters, is not large enough, then the process stops with no ellipse detected nor matched.
- In a practical settings, when a given accumulator peak has been obtained, we look back for all the evidences that voted for this peak. These evidences are then removed from the initial list of evidences and we repeat the process from the beginning until no ellipses are detected anymore and the process stops.

Obviously, managing a $5D$ accumulator may be very time and memory consuming in practice. Fortunately, solutions exist to reduce the dimension of the accumulator.

3.3 First Step: Reducing the Problem Dimension

Projection of the dual parameters to (C, C')

In the literature, without always explicitly noting it, authors reduce the dimension of the problem by considering a projection (of the process explained in the previous section) to a lower dimension space. For example, some geometric observations lead to the solution chosen by all authors that consists in accumulating the line l joining the mid-point M of the edge points P and Q from one side and the intersection point O of the two tangents q and p from the other side (see Figure 1). This line is always a diameter of the unknown ellipse and passes through the ellipse center. The accumulation of this line in a plane whose dimensions are similar to those of the image highlights the possible ellipse centers.

In the framework of the pencil of bi-tangent conics, this method is equivalent to the projection of the line representing the pencil in the $5D$ space to a line in the $2D$ space. Figure (2) illustrates the process for the dual approach; the blue line in the $5D$ space (represented in $3D$ in the figure) represents the locus of the parameters (C, C', D, D', D'') of the conics in the pencil defined by a specific pair of edge points and the blue line in the plane (C, C') represents its projection.

From equation (14), we see that a degenerate case occurs when the gradients are parallel. In this situation, the $5D$ line is exactly orthogonal to the (C, C') plane and its projection reduces to a point. This means that, when the gradients are parallel, the position of the ellipse center is known perfectly.

Projection of the direct parameters to (d, d')

However, there are other possibilities to reduce the dimensions of the problem. For instance, we could also project the line in the $5D$ space (d, d', c, c', c'') into a line in the $2D$ space (d, d'). Indeed, from equation (3), we know that these (d, d') parameters are equivalent to the eccentricity e (aspect ratio) and orientation θ of the ellipse. An interesting property of these parameters is that they are always comprised within the unit circle of their plane. The domain of the accumulator is then well defined and bounded. Also, based on equation (13), it is possible to demonstrate that, in the case of this new projection, there is no degenerated case and the $5D$ line is never exactly orthogonal to the (d, d') plane.

Accumulation in $2D$

Regardless of the chosen (direct or dual) $2D$ projection, the first step of the algorithm is to iterate on all pairs of edge points which verify the conditions expressed in equations (11), and to draw a line in the corresponding plane ($2D$ accumulator). In Figure (2), the gray lines in the plane (C, C') are instances, in the dual framework, of these $2D$ lines (diameters OM) for several pairs of edge points.

When the accumulation or voting process is complete, we look for the maximum value (or the most prominent cluster). If this value (or the sum of the cluster values) is not large enough, we stop the process with no ellipse detected. Otherwise, the position (a $2D$ index) of this peak (or cluster center) is chosen as the estimated value (Δ, Δ') of the two ellipse parameters (the big red point in Figure (2)).

Edge points selection

In the dual framework, when the detected center is close, in the $2D$ space (C, C'), to a line (diameter OM) related to a given pair of edge points, we consider that this pair of edge points has voted for the detected center and we select this pair for the second step of the algorithm. In Figure (2), the blue line of the plane (C, C') is a $2D$ line which is related to a specific pair of edge points. It passes near the detected center and is then selected as having voted for this center.

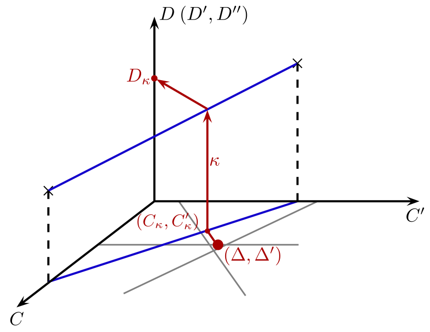

Figure 2: Back-projection of the estimated center (Δ, Δ') into a $3D$ space (illustration of the back-projection into a $5D$ space)

3.4 Second Step: Back-projection

State of the Art Methods

In the literature, authors switch to other methods or algorithms to find the remaining parameters. Most of the time, they use the information about the center location to obtain the equation of the centered ellipse. Then, all the edge points (and their gradients) that previously voted for the detected center provide now two conditions for the three remaining parameters. So, by a similar process of line accumulation in a three dimension space, they obtain an estimation for the remaining parameters.

The Proposed Algorithm

In the framework of the pencil of bi-tangent conics, the three remaining parameters can advantageously be computed by back-projecting the points obtained during the $2D$ accumulation process onto the line in the full $5D$ space.

Index of the closest ellipse

In Figure 2, the detected center (Δ, Δ') (big red point) is orthogonally projected onto the $2D$ line (blue line in the (C, C') plane) related to a specific pair of edge points. As a point of the line, this orthogonal projection should be of the form (C_κ, C'_κ) (given by the two first equations in (12) and (14)), where κ is a specific value for the index k. In the following, κ will be named the index of the closest conic in the pencil defined by this specific pair of edge points.

Likewise, for the direct approach, the same reasoning holds for finding the index of the couple of parameters that is the closest to the estimated values in the (d, d') plane.

However, there is a major difference between the two approaches. In the dual approach, when the gradients are parallel, the $5D$ line is orthogonal to the plane (C, C'). So, in this situation, it is not possible to deduce the value of κ from the estimated center and, for this step, we must discard the pairs of edge points whose gradients are parallel. In the direct approach, there is no such situation; the $5D$ line is never orthogonal to the plane (d, d') and all pairs of edge points may be used for the back-projection step.

The three remaining parameters

In order to compute the three remaining ellipse parameters (D, D', D''), we use, in equation (14), this index κ of the closest ellipse whose the parameters (C_κ, C'_κ) are the closest to the estimated center (Δ, Δ'). This means that a pair of edge points, which voted for a detected center, also provides (or votes for) one $3D$ vector of the remaining parameters. In Figure (2), a point (C_κ, C'_κ) on the $2D$ line corresponds to one $5D$ point $(C_\kappa, C'_\kappa, D_\kappa, D'_\kappa, D''_\kappa)$.

The last step is to iterate on all the pairs of edge points supporting the estimated center and to fill a $3D$ histogram (or three $1D$ histograms) with the provided vectors of the remaining parameters. If the peak of (or the most prominent cluster in) the histogram is not large enough, no ellipse is detected nor matched. Otherwise, the $3D$ vector of indices of the peak (or cluster center) is chosen as the estimated value for the three remaining parameters.

Likewise, the same reasoning also holds in the direct approach for finding the three remaining parameters (c, c', c'').

Finally, we remove from the list of pairs of edge points those supporting the chosen vector of remaining parameters and we start again the whole process until no ellipses are detected anymore.

4 Results

To illustrate the feasibility of the methods introduced here, we present some preliminary comparison results between our dual algorithm and the State-of-the-Art (abbreviated SoA in the following) algorithm outlined in Section 3.4. Both algorithms have been implemented with the same list of edge points, the same center detection first step, the same accumulator cell size and without any post-processing. Obviously, in a practical settings, both algorithms could benefit from additional processing step, but we are interested here in the bare properties of the methods.

We generate 100 synthetic images (450 by 600 pixels) containing one random black ellipse on a white background (see Figure 3) and we add salt and pepper noise and a randomly (Gaussian) textured image.

We compute the Hausdorff distance between the detected and the true ellipse for each test image and we display their statistics in Table 1. We also display in Table 2 the number of detected ellipses whose Hausdorff distance to the true ellipse is respectively less than 2, 5, 10 and 15 pixels.

	Dual	SoA
Mean	2.3	6.1
Std Dev	1.5	8.4
Min	0.6	0.8
Max	10.6	59

Table 1: Statistics of the Hausdorff distance between the detected and true ellipse

Threshold	Dual	SoA
2	56	30
5	94	67
10	99	85
15	100	90

Table 2: Number of detected ellipses whose Hausdorff distance to the true ellipse is less than the threshold given in the first column

In this experiment, we observe that the dual approach provides a better estimation (in terms of Hausdorff distance) of the ellipse than our implementation of the SoA approach. In addition, if we consider that an ellipse is not (correctly) detected when its Hausdorff distance to the true ellipse is larger than 10, from the Table 2, we see that, in our experiment, 1% (resp. 15%) of the ellipse were not detected by the dual (resp. SoA) approach.

Figure 3 shows a typical example of one test image (left) with the true ellipse delineated in red. The background of the middle and the right images represents the same edge points (in white) extracted from the left image. In the middle (resp. right) image, the green curve is the detected ellipse and the red points are the selected edge points, both by our implementation of the SoA (resp. dual) algorithm. The selected edge points (in red) are chosen as being closer than 10 pixels to the ellipse detected and may be used as the input of an additional ellipse fitting step, if necessary. On this example, the Hausdorff distance between the detected and the true ellipse is 11.8 for the SoA algorithm and 3.4 for the dual approach. We then see the impact of a poor precision for the ellipse parameter estimation on the capacity of the algorithm to correctly extract the edge points belonging to an ellipse.

Figure 3: Image and edge points

5 Conclusions

In this paper, we show that the concept of pencil of bi-tangent conics, either in its direct or dual form, is an adequate framework to discuss the theoretical aspects of the EHT based on pairs of edge points and their associated gradients. The usual process of splitting the full $5D$ problems into two (or three sometimes) steps of reduced dimension is well modeled by the notion of projection of the problem to a $2D$ sub-space followed by a back-projection of the results into the full $5D$ space.

The algebraic coherence allows us to unify both parts of the EHT into a single framework, highlights the eligibility conditions of the pairs of edge points (to be possible evidence of the presence of an ellipse) and manages clearly special cases for the pairs of edges points or their gradients.

The dual form of the framework models the first step of the state-of-the-art methods of finding the ellipse center but extracts differently the remaining parameters. Our preliminary results suggest that this new approach could have a better accuracy in the estimation of the ellipse parameters.

The direct form of the framework is, as far as we now, a new approach to reduce the dimension and solve the problem. This direct approach first looks for a couple of ellipse parameters similar to the eccentricity and the orientation. Theoretically, both approaches have benefits and drawbacks; the choice between them will be a matter of practical considerations.

We hope that this presentation of the direct and dual equations of the pencil of bi-tangent conics for the EHT will pave the way for new approaches.

References

[1] K. Kanatani, Y. Sugaya, and Y. Kanazawa. *Ellipse Fitting for Computer Vision: Implementation and Applications*. Synthesis Lectures on Computer Vision. Morgan & Claypool, 2016.

[2] R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, January 1972.

[3] P. Mukhopadhyay and B. Chaudhuri. A survey of hough transform. *Pattern Recognition*, 48(3):993–1010, March 2015.

[4] S. Tsuji and F. Matsumoto. Detection of ellipses by a modified Hough transformation. *IEEE Transactions on Computers*, C-27(8):777–781, August 1978.

[5] H. K. Yuen, J. Illingworth, and J. Kittler. Ellipse detection using the Hough transform. In *Proceedings of the Alvey Vision Conference*, pages 41.1–41.8. Alvey Vision Club, 1988.

[6] N. Guil and E. L. Zapata. Lower order circle and ellipse Hough transform. *Pattern Recognition*, 30(10):1729–1744, October 1997.

[7] H. K. Muammar and M. Nixon. Tristage Hough transform for multiple ellipse extraction. *IEE Proceedings E, Computers and Digital Techniques*, 138(1):27–35, 1991.

[8] J. H. Yoo and I. K. Sethi. An ellipse detection method from the polar and pole definition of conics. *Pattern Recognition*, 26(2):307–315, February 1993.

[9] N. Bennett, R. Burridge, and N. Saito. A method to detect and characterize ellipses using the Hough transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7):652–657, July 1999.

[10] A. B. Forbes. Fitting an Ellipse to Data. NPL Report DITC 95/87, National Physical Laboratory, 1987.

[11] V. F. Leavers. The dynamic generalized Hough transform: Its relationship to the probabilistic Hough transforms and an application to the concurrent detection of circles and ellipses. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 56(3):381–398, November 1992.

[12] W. Lu, J. Yu, and J. Tan. Direct inverse randomized Hough transform for incomplete ellipse detection in noisy images. *Journal of Pattern Recognition Research*, 1:13–24, 2014.