Representations that uniquely characterize images modulo translation, rotation, and scaling

Robert D. Brandt a,1, Feng Lin b,∗

a Intelligent Devices, Inc., 465 Whittier Avenue, Glen Ellyn, IL 60137, USA
b Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA

Received 15 June 1995; revised 25 March 1996

Abstract

Representations that uniquely characterize images modulo the planar geometric similarity transformations are presented. Such representations are invariant with respect to translation, rotation, scaling, and their combination. We will first discuss basic invariants, that is, the invariants for translation, rotation, and scaling only. We will then discuss hybrid invariants, that is, invariants with respect to combinations of translation, rotation, and scaling.

Keywords: Invariant; Translation; Rotation; Scaling

1. Introduction

In this paper, we consider the problem of recognizing isolated objects from their images, regardless of the position, orientation, and scale at which they are represented. Such a recognition process is said to be invariant with respect to the (planar) geometric similarity transformations: translation, rotation, and scaling. The most natural representation of an image is as a non-negative function of the Cartesian coordinates of the image plane. The value of the function usually represents the intensity of light reflected to the eye. We refer to images represented in this way as raw images. In order to directly verify that two raw images are of the same object, comparison of the two images must be made for every possible similarity transformation until a match is determined. Such comparisons could be performed either by using the cross-correlation of the images or by using a metric, such as the $L_2$ metric (Ballard and Brown, 1982; Gonzalez and Woods, 1992; Rogers et al., 1990).

The number of computations required to perform a digital approximation of this idealized matching process depends on the resolution (sample density) of the sampled images and the accuracy required to justify the confirmation of a match. Match of the four (quantized) transformation parameters (two for translation, one for rotation, and one for scaling) contributes to the number of calculations that must be performed. In applications with a large number of candidate objects, such brute force comparisons are unfeasible.

∗ Corresponding author. E-mail: flin@ece.eng.wayne.edu.
1 Currently visiting the Beckman Institute, University of Illinois, Urbana, IL 61801, USA.
In this paper we examine alternative representations which allow comparisons to be made more efficiently. The basic idea is to construct representations that are invariant with respect to some or all of the similarity transformations. In this way the complexity of comparisons is reduced by eliminating the necessity of comparing the images for every possible relative transformation. The trade-off is the extra computation required to calculate the invariant representations from the raw images.

Only representations that retain all information except for position, orientation, and scale are considered. We refer to such representations as complete representations to distinguish them from representations that do not uniquely characterize images up to position, orientation, and scale. Incomplete representations can be useful for certain applications. For example a very useful feature vector for the recognition of machine-printed characters is composed of the areas of convex regions formed by the curving of the lines comprising the character. These features are invariant, but retain only a small portion of the original information, which happens to be sufficient for the recognition of machine-printed characters (Brandt, 1992). Feature domain representations and other incomplete representations are not the focus of this paper.

The most obvious complete invariant image representations are derived by image normalization, the process by which images are converted to a canonical form which is invariant with respect to some or all of the similarity transformations. One such canonical form is the image whose centroid has been translated to the center (origin) of the image plane, and whose integral (mass) has been made 1 (or any other chosen constant) by scaling the image. The normalized forms of two images will match for some orientation if the two images are indeed of the same object. Thus, the search need only be performed over possible orientations, and searches over positions and scales will have been avoided. Reduced search time is replaced with the extra processing required to convert the raw image into normalized form. In some applications, it is possible to normalize the orientation. Normalized representations are referred to as spatial domain invariants.

Frequency domain invariants are derived from the Fourier transforms of images. For example, it is easy to show that the magnitude spectrum of an image is invariant with respect to translation. The representation is incomplete, because meaningful relative phase information is discarded along with the undesired linear component of the phase. The Taylor invariant (Lin and Brandt, 1993) and the Hessian invariant derived in this paper solve this problem by eliminating only the linear component of the phase. Similarly, the magnitude of the Fourier–Mellin transform is invariant with respect to rotation and scaling, but is incomplete. Orientation and scale analogs of the Taylor and Hessian invariants provide the relative phase information that is missing from the Fourier–Mellin magnitude. Related work can be found in (Mundy and Zisserman, 1992; Reiss, 1993; Castro and Morandi, 1987; Jacobson and Wechsler, 1984; Sheng, 1989).

Another approach is to represent the images with moment domain invariants. These often-studied invariants establish a bridge between spatial domain invariants and frequency domain invariants (Armstrong et al., 1990; Belkasim et al., 1991; Dudani et al., 1977; Hsia, 1981; Hu, 1962; Maitra, 1979).

In this paper, we will study invariants in the frequency domain. We will start with the basic invariants, these are invariant under translation, rotation, or scaling only. Our main results are on hybrid invariants, that is, invariants with respect to combinations of translation, rotation, and scaling. To find hybrid invariants, we need to introduce the properties of rotational symmetry and reciprocal scaling. These properties are needed in combining two invariants. They guarantee that the composition of two invariants, each with respect to a group, is also an invariant with respect to both groups. Although there are many ways to combine the various basic invariants, we will concentrate on Euclidean invariants, translation and scaling invariants, and similarity invariants. Since our invariants are obtained in the frequency domain without spatial differentiation, they are global invariants.
2. Definitions and notations

In this section, we give some basic definitions and clarify our basic notation. We will first discuss important definitions and basic properties of images and their raw, Fourier, and Fourier–Mellin representations. We will also discuss coordinate transformations and invariant operators that will be used throughout the paper. Readers already familiar with these definitions and style of notation may prefer to skip this section and refer to it as needed.

2.1. Image representation

We say that a function \( f(x, y) \) has finite energy if

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)^2 \, dx \, dy < \infty.
\]

We say that a function \( f(x, y) \) has a bounded support if there exists a constant \( B \) such that for all \( x \) and \( y \):

\[
|x| > B \quad \text{and} \quad |y| > B,
\]

\[
f(x, y) = 0.
\]

Definition 1 (Raw image). A raw image is a piecewise smooth (twice differentiable), non-negative function with finite energy and bounded support in the image plane.

In this paper we will consistently use the symbol, \( f \), to represent an arbitrary raw image. The value, \( f(x, y) \geq 0 \), typically represents the intensity of light reaching a particular point, \((x, y)\), of an imaging device. The images we have in mind are of single isolated objects on a constant background.

Definition 2 (Image representation). An image representation, \( J \), is an operator on the set of raw images.

The raw image itself can be considered as a representation (by way of the identity operator).

Remark 1 (Operator notation). In this paper we use the following notation: if \( O \) is an operator, then the notation, \( O(x, y)[f] \) (or simply \( O(x, y) \) if \( f \) is understood), is interpreted as \((O[f])(x, y)\). Thus, the notation, \( O(x, y) \), represents a functional (an operator with range in the set of real numbers).

The Fourier representation is the primary frequency domain representation from which other frequency domain representations are derived. It is defined as follows.

Definition 3 (Fourier representation). The Fourier representation, \( F \), of a raw image, \( f \), is given by

\[
F(\omega_x, \omega_y)[f] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(x\omega_x + y\omega_y)} f(x, y) \, dx \, dy.
\]

Since the Fourier representation is a complex function, we often study its magnitude and phase separately.

Definition 4 (Magnitude representation). The Fourier magnitude representation, \( A \), is the absolute value of the Fourier representation. Symbolically,

\[
A(\omega_x, \omega_y) = \text{abs} \, F(\omega_x, \omega_y).
\]
**Definition 5 (Phase representation).** The Fourier phase representation, $\Phi$, is the argument of the Fourier representation. Symbolically,

$$\Phi(\omega_x, \omega_y) = \arg F(\omega_x, \omega_y) \mod 2\pi.$$  

Strictly speaking, the Fourier phase representation is defined modulo $2\pi$; however, in certain cases of interest it is possible to make a well-defined choice of branch at each point.

To discuss rotation and scaling invariants, it is more convenient to have the raw image represented in the logarithmic-polar coordinates. Also the Fourier–Mellin representation will be used if the raw image is given in the logarithmic-polar coordinates.

**Definition 6 (Logarithmic-polar coordinates).** The logarithmic-polar transformation is given by

$$LP(\mu, \xi)[f] = \begin{cases} 
  f(e^\mu \cos \xi, e^\mu \sin \xi) & \xi \in [0, 2\pi), \\
  0 & \xi \notin [0, 2\pi). 
\end{cases}$$

We can take the Fourier–Mellin transform of the above representation with respect to $\mu, \xi$, or both to obtain the following representation.

**Definition 7 (Fourier–Mellin representation).** The Fourier transform of $LP(\mu, \xi)[f]$ with respect to $\mu$ is given by

$$F_S(\omega, \xi) = \int_{-\infty}^{\infty} e^{i\omega \mu} LP(\mu, \xi) \, d\mu.$$  

The Fourier series of $LP(\mu, \xi)[f]$ with respect to $\xi$ is given by

$$F_R(\mu, k) = \int_{0}^{2\pi} e^{ik \xi} LP(\mu, \xi) \, d\xi.$$  

The Fourier–Mellin representation, $F_m$, of a raw image, $f$, is given by

$$F_m(\omega, k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\omega \mu + k \xi)} LP(\mu, \xi) \, d\xi \, d\mu.$$  

We now give the following definitions because it is often useful to study the magnitude and phase of the Fourier–Mellin representation separately.

**Definition 8 (Magnitude representation).** The magnitudes of $F_S(\omega, \xi)$, $F_R(\mu, k)$ and $F_m(\omega, k)$ are denoted by

$$A_S(\omega, \xi) = \text{abs} F_S(\omega, \xi); \quad A_R(\mu, k) = \text{abs} F_R(\mu, k); \quad A_m(\omega, k) = \text{abs} F_m(\omega, k).$$  

**Definition 9 (Phase representation).** The phases of $F_S(\omega, \xi)$, $F_R(\mu, k)$ and $F_m(\omega, k)$ are denoted by

$$\Phi_S(\omega, \xi) = \arg F_S(\omega, \xi) \mod 2\pi; \quad \Phi_R(\mu, k) = \arg F_R(\mu, k) \mod 2\pi; \quad \Phi_m(\omega, k) = \arg F_m(\omega, k) \mod 2\pi.$$
2.2. Coordinate transformation

As stated before, we are interested in translation, rotation and scaling transformations. They are formally defined as follows.

**Definition 10 (Translation transformation).** The (commutative) group of translation transformations, \( T = \{T_{\alpha,\beta}\} \), where \( \alpha \) and \( \beta \) range over real numbers, is given by

\[
T_{\alpha,\beta}(x, y)[f] = f(x + \alpha, y + \beta).
\]

**Definition 11 (Scaling transformation).** The (commutative) group of scaling transformations (homotheties), \( S = \{S_{\rho}\} \), where \( \rho \) ranges over positive real numbers, is given by

\[
S_{\rho}(x, y)[f] = f(\rho x, \rho y).
\]

**Definition 12 (Rotation transformation).** The (commutative) group of rotation transformations, \( R = \{R_{\theta}\} \), where \( \theta \) ranges over real numbers, is given by

\[
R_{\theta}(x, y)[f] = f(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta).
\]

**Definition 13 (Rotation and scaling transformation).** The (commutative) group of rotation and scaling transformations, \( RS = \{RS_{\theta,\rho}\} \), is given by

\[
RS_{\theta,\rho} = R_{\theta} \circ S_{\rho} = S_{\rho} \circ R_{\theta}.
\]

2.3. Invariant operators

Our goal is to find operators that are invariant under various transformations described above. Formally, invariant operators are defined as follows.

**Definition 14 (Invariant).** An operator, \( O \), is an invariant with respect to a group of transformations \( Q \) if

\[
(\forall Q \in Q)O \circ Q = O.
\]

All the invariants discussed in this paper are complete invariants as described below.

**Definition 15 (Complete invariant).** An operator, \( O \), is a complete invariant with respect to a group of transformations \( Q \) if it is invariant with respect to \( Q \) and

\[
O[f_1] = O[f_2] \Rightarrow (\exists Q \in Q)f_1 = Q[f_2].
\]

Only complete invariants capture all the information of an image modulo the given transformation. Most of invariants discussed so far in literature are not complete.

2.4. Composition of transformations

In many applications, images are transformed not under just one single transformation as discussed above, but rather under combination of several transformations. So, we study the following compositions of transformations.

**Definition 16 (Euclidean transformation).** The (non-commutative) group of Euclidean (translation and rotation) transformations, \( RT = \{RT_{\theta,\alpha,\beta}\} \) is given by
\[ RT_{\theta, \alpha, \beta} = R_{\theta} \circ T_{\alpha, \beta}. \]

Note \( T_{\alpha, \beta} \circ R_{\theta} \neq R_{\theta} \circ T_{\alpha, \beta} \), but rather, \( T_{\alpha, \beta} \circ R_{\theta} = R_{\theta} \circ T_{\alpha \cos \theta - \beta \sin \theta, \alpha \sin \theta + \beta \cos \theta} \). However, if an operator is a (complete) invariant with respect to \( R_{\theta} \circ T_{\alpha, \beta} \), it is also a (complete) invariant with respect to \( T_{\alpha, \beta} \circ R_{\theta} \). Therefore, we only need to discuss the group \( RT \).

**Definition 17** (*Translation and scaling transformation*). The (non-commutative) group of translation and scaling transformations, \( ST = \{ ST_{p, \alpha, \beta} \} \) is given by

\[ ST_{p, \alpha, \beta} = S_{p} \circ T_{\alpha, \beta}. \]

Note \( T_{\alpha, \beta} \circ S_{p} \neq S_{p} \circ T_{\alpha, \beta} \), but rather, \( T_{\alpha, \beta} \circ S_{p} = S_{p} \circ T_{p \alpha \beta} \). However, if an operator is a (complete) invariant with respect to \( S_{p} \circ T_{\alpha, \beta} \), it is also a (complete) invariant with respect to \( T_{\alpha, \beta} \circ S_{p} \). Therefore, we only need to discuss the group \( ST \).

**Definition 18** (*Similarity transformation*). The (non-commutative) group of similarity (translation, rotation, and scaling transformations, \( RST = \{ RST_{\theta, p, \alpha, \beta} \} \) is given by

\[ RST_{\theta, p, \alpha, \beta} = RS_{\theta, p} \circ T_{\alpha, \beta}. \]

Note \( T_{\alpha, \beta} \circ RS_{\theta, p} \neq RS_{\theta, p} \circ T_{\alpha, \beta} \), but rather, \( T_{\alpha, \beta} \circ RS_{\theta, p} = RS_{\theta, p} \circ T_{p \alpha \cos \theta - \beta \sin \theta, p \alpha \sin \theta + \beta \cos \theta} \). However, if an operator is a (complete) invariant with respect to \( RS_{\theta, p} \circ T_{\alpha, \beta} \), it is also a (complete) invariant with respect to \( T_{\alpha, \beta} \circ RS_{\theta, p} \). Therefore, we only need to discuss the group \( RST \).

### 2.5. Operator symmetry

Let us consider an operator \( O \) that maps an image from the spatial domain to the frequency domain. The following symmetry properties of the operator are important in constructing hybrid invariants (invariants with respect to compositions of transformations).

**Definition 19** (*Rotational symmetry*). An operator, \( O \), is rotationally symmetric if for all \( R_{\theta} \in \mathbb{R} \),

\[ O \circ R_{\theta} = R_{\theta} \circ O. \]

In words, \( O \) is rotationally symmetric if rotation in the spatial domain corresponds to rotation in the frequency domain.

**Definition 20** (*Reciprocal scaling*). An operator, \( O \), is reciprocally scaled if for all \( S_{p} \in \mathbb{S} \),

\[ O \circ S_{p} = S_{1/p} \circ O. \]

In words, \( O \) is reciprocally scaled if expansion in the spatial domain corresponds to contraction in the frequency domain.

**Definition 21** (*Rotational symmetry and reciprocal scaling*). An operator, \( O \), is rotationally symmetric and reciprocally scaled if for all \( RS_{\theta, p} \in \mathbb{R} \),

\[ O \circ RS_{\theta, p} = RS_{\theta, 1/p} \circ O. \]
3. Translation invariants

The following two complete translation invariants, the Taylor and Hessian invariants, have been derived in (Lin, 1993). Both are frequency domain invariants. In the theorems, $D_{\omega_i}$ denotes the derivative with respect to $\omega_i$, etc.

**Theorem 1** (Taylor invariant). The representation given by

$$F_c(\omega_x, \omega_y) = e^{-i(a_{\omega_x} + b_{\omega_y})} F(\omega_x, \omega_y),$$

where

$$a = D_{\omega_x}(0, 0) \circ \Phi[f], \quad b = D_{\omega_y}(0, 0) \circ \Phi[f],$$

is a complete translation invariant.

**Theorem 2** (Hessian invariant). The representation given by

$$F_H(\omega_x, \omega_y) = \begin{bmatrix} A(\omega_x, \omega_y) & 0 & 0 \\ 0 & \Phi_{\omega_x, \omega_x}(\omega_x, \omega_y) & \Phi_{\omega_x, \omega_y}(\omega_x, \omega_y) \\ 0 & \Phi_{\omega_y, \omega_x}(\omega_x, \omega_y) & \Phi_{\omega_y, \omega_y}(\omega_x, \omega_y) \end{bmatrix},$$

where

$$\Phi_{\omega_x, \omega_x} = D_{\omega_x} \circ D_{\omega_x} \circ \Phi, \quad \Phi_{\omega_y, \omega_y} = D_{\omega_y} \circ D_{\omega_y} \circ \Phi,$$

$$\Phi_{\omega_x, \omega_y} = D_{\omega_x} \circ D_{\omega_y} \circ \Phi, \quad \Phi_{\omega_y, \omega_x} = D_{\omega_y} \circ D_{\omega_x} \circ \Phi,$$

is a complete translation invariant.

By the assumption of finite energy and bounded support, $\Phi_{\omega_x, \omega_x} = \Phi_{\omega_y, \omega_y}$. Therefore the Hessian invariant consists of four distinct elements. We can recover the complete image modulo translation from its Taylor or Hessian invariants. For the Taylor invariant, this is straightforward: we just need to take the inverse Fourier transform of $F_c(\omega_x, \omega_y)$. For the Hessian invariant, we can take the inverse Fourier transform of

$$A(\omega_x, \omega_y) e^{i\Phi'(\omega_x, \omega_y)},$$

where $\Phi'(\omega_x, \omega_y)$ is given by

$$\int_{0}^{\omega_x} \int_{0}^{\omega_y} \phi_{\omega_x, \omega_y}(s, t) \, ds \, dt + \int_{0}^{\omega_x} \phi_{\omega_x, \omega_y}(s, 0) \, ds + \int_{0}^{\omega_y} \phi_{\omega_y, \omega_x}(0, t) \, dt.$$

We have implemented the reconstruction formula for Hessian invariant in MATLAB and tested several images. One typical result is shown in Fig. 1. In the figure, (a) is the original image. We take its Fourier transform, derive its Hessian invariant, and then reconstruct the image using the above formula. The result is shown in (b).

4. Rotation and scaling invariants

Rotation and scaling invariants can be derived in a manner similar to translation invariants, using the logarithmic-polar coordinates and Fourier–Mellin transforms. They are presented in the following theorems, whose proofs are similar to that of Theorem 2.
Theorem 3 (Mellin-S invariant). The representation given by

\[ M_\mathcal{S}(\omega, \xi) = \begin{bmatrix} A_\mathcal{S}(\omega, \xi) & 0 & 0 \\ 0 & \Phi_{\text{Sovo}}(\omega, \xi) & \Phi_{\text{Sov}}(\omega, \xi) \\ 0 & \Phi_{\text{Sov}}(\omega, \xi) & \Phi_{\text{Svov}}(\omega, \xi) \end{bmatrix}, \]

where

\[ \Phi_{\text{Sovo}}(\omega, \xi) = D_\omega(\omega, \xi) \circ D_\omega \circ \Phi_\mathcal{S}, \]

\[ \Phi_{\text{Sov}}(\omega, \xi) = D_\omega(\omega, \xi) \circ D_\xi \circ \Phi_\mathcal{S}, \]

\[ \Phi_{\text{Svov}}(\omega, \xi) = D_\xi(\omega, \xi) \circ D_\xi \circ \Phi_\mathcal{S}, \]

is a complete scaling invariant.

Theorem 4 (Mellin-R invariant). The representation given by
\[ M_{R}(\mu, k) = \begin{bmatrix}
A_{R}(\mu, k) & 0 & 0 \\
0 & \Phi_{R\mu}(\mu, k) & \Phi_{Rk}(\mu, k) \\
0 & \Phi_{R\mu}(\mu, k) & \Phi_{Rk}(\mu, k)
\end{bmatrix}, \]

where

\[ \Phi_{R\mu}(\mu, k) = D_{\mu}(\mu, k) \circ D_{\mu} \circ \Phi_{R}, \]
\[ \Phi_{Rk}(\mu, k) = D_{\mu}(\mu, k) \circ \Phi_{R} - D_{\mu}(\mu, k - 1) \circ \Phi_{R}, \]
\[ \Phi_{Rkk}(\mu, k) = \Phi_{R}(\mu, k) - 2\Phi_{R}(\mu, k - 1) + \Phi_{R}(\mu, k - 2), \]

is a complete rotation invariant.

**Theorem 5 (Mellin-RS invariant).** The representation given by

\[ M_{RS}(\omega, k) = \begin{bmatrix}
A_{m}(\omega, k) & 0 & 0 \\
0 & \Phi_{m\mu}(\omega, k) & \Phi_{mk}(\omega, k) \\
0 & \Phi_{m\mu}(\omega, k) & \Phi_{mk}(\omega, k)
\end{bmatrix}, \]

where

\[ \Phi_{m\mu}(\omega, k) = D_{\omega}(\omega, k) \circ D_{\omega} \circ \Phi_{m}, \]
\[ \Phi_{mk}(\omega, k) = D_{\omega}(\omega, k) \circ \Phi_{m} - D_{\omega}(\omega, k - 1) \circ \Phi_{m}, \]
\[ \Phi_{mkk}(\omega, k) = \Phi_{m}(\omega, k) - 2\Phi_{m}(\omega, k - 1) + \Phi_{m}(\omega, k - 2), \]

is a complete rotation and scaling invariant.

Note that the Mellin-R and Mellin-RS invariants are discrete in \( k \), because of the periodicity of \( \theta \).

The complete image modulo rotation and/or scaling can be recovered from the above complete invariants in a way similar to that of the translation invariants discussed in the previous section.

**5. Construction of hybrid invariants**

In order to simplify notations in constructing hybrid invariants, let us first introduce the following notations.

**Remark 2 (Simplified notations).** Note that the Hessian invariant (a 3 \( \times \) 3 matrix) is uniquely determined by 4 of its elements. Therefore, in the rest of the paper, we denote the Hessian invariant as a vector to simplify the notation:

\[ F_{H}(\omega_{x}, \omega_{y}) = \begin{bmatrix}
A(\omega_{x}, \omega_{y}) \\
\Phi_{\omega_{x}, \omega_{x}}(\omega_{x}, \omega_{y}) \\
\Phi_{\omega_{x}, \omega_{y}}(\omega_{x}, \omega_{y}) \\
\Phi_{\omega_{x}, \omega_{x}}(\omega_{x}, \omega_{y})
\end{bmatrix}. \]

Similarly,
Remark 3 (Composition of operators). We use the following convention when combining two operators. If $O_1$ is multi-valued,

$$O_1 = \begin{bmatrix} O_{11} \\ \vdots \\ O_{1n} \end{bmatrix},$$

and $O_2$ is single-valued, then

$$O_1 \circ O_2 = \begin{bmatrix} O_{11} \circ O_{2} \\ \vdots \\ O_{1n} \circ O_{2} \end{bmatrix}.$$

If $O_3$ is multi-valued,

$$O_3 = \begin{bmatrix} O_{31} \\ \vdots \\ O_{3m} \end{bmatrix},$$

then

$$O_3 \circ O_1 = \begin{bmatrix} O_{31} \circ O_{11} & \cdots & O_{31} \circ O_{1n} \\ \vdots & \ddots & \vdots \\ O_{3m} \circ O_{11} & \cdots & O_{3m} \circ O_{1n} \end{bmatrix}.$$

Let us first describe our approach to construct hybrid invariants. Given two operators $O_1$ and $O_2$ that are invariant with respect to two groups of transformations $Q_1$ and $Q_2$, respectively. That is, for all $Q_1 \in Q_1$, $Q_2 \in Q_2$, the diagrams in Fig. 2 commute.

Now, we hope that the operator $O = O_1 \circ O_2$ is invariant with respect to the composition of the two transformations $Q = Q_2 \circ Q_1$. In other words, we hope that the diagram in Fig. 3 commutes.

This will be the case (that is, the dashed lines become solid) if $Q_2 \circ Q_1 = O_1 \circ Q_2$. To ensure that $Q_2 \circ O_1 = O_1 \circ Q_2$, we need the properties of rotational symmetry and reciprocal scaling as shown in the following theorems.

Theorem 6 (Euclidean invariants). Suppose an operator $O_1$ is invariant with respect to $T$ and is rotationally symmetric, and another operator $O_2$ is invariant with respect to $R$. If the domain of $O_2$ contains the range of $O_1$, then the hybrid operator $O = O_2 \circ O_1$ is invariant with respect to $RT$. The invariant $O$ is a complete invariant if $O_1$ and $O_2$ are complete invariants.

Proof. $O$ is invariant with respect to $RT$ because for all $RT_{\theta,\alpha,\beta} \in RT$,

$$O \circ RT_{\theta,\alpha,\beta} = O_2 \circ O_1 \circ R_{\theta} \circ T_{\alpha,\beta} = O_2 \circ R_{\theta} \circ O_1 \circ T_{\alpha,\beta} = O_2 \circ O_1 = O.$$

On the other hand, if $O_1$ and $O_2$ are complete invariants,
\[ O[f_1] = O[f_2] \implies O_2 \circ (O_1[f_1]) = O_2 \circ (O_1[f_2]) \]
\[ \implies (\exists \theta) O_1[f_1] = R_\theta \circ O_1[f_2] \]
\[ \implies (\exists \theta) O_1[f_1] = O_1 \circ R_\theta[f_2] \]
\[ \implies (\exists \theta)(\exists \alpha, \beta) f_1 = T_{\alpha, \beta} \circ R_\theta[f_2] \]
\[ \implies (\exists \theta)(\exists \alpha, \beta) f_1 = R_\theta \circ T_{\alpha \cos \theta - \beta \sin \theta, \alpha \sin \theta + \beta \cos \theta}[f_2] \]
\[ \implies (\exists \theta)(\exists \alpha, \beta) f_1 = R_{T_{\alpha \cos \theta - \beta \sin \theta, \alpha \sin \theta + \beta \cos \theta}}[f_2]. \]

Therefore, \( O \) is a complete invariant. \( \Box \)

**Theorem 7** (Translation and scaling invariants). Suppose an operator \( O_1 \) is invariant with respect to \( T \) and is reciprocally scaled, and another operator \( O_2 \) is invariant with respect to \( S \). If the domain of \( O_2 \) contains the range of \( O_1 \), then the hybrid operator \( O = O_2 \circ O_1 \) is invariant with respect to \( ST \). The invariant \( O \) is a complete invariant if \( O_1 \) and \( O_2 \) are complete invariants.

**Proof.** The proof is similar to the proof of Theorem 6. \( \Box \)

**Theorem 8** (Similarity invariants). Suppose an operator \( O_1 \) is invariant with respect to \( T \), and is rotationally symmetric and reciprocally scaled. Also suppose another operator \( O_2 \) is invariant with respect to \( RS \). If the domain of \( O_2 \) contains the range of \( O_1 \), then the hybrid operator \( O = O_2 \circ O_1 \) is invariant with respect to \( RST \). The invariant \( O \) is a complete invariant if \( O_1 \) and \( O_2 \) are complete invariants.

**Proof.** The proof is similar to the proof of Theorem 6. \( \Box \)

With these three theorems, we can construct various hybrid invariants in the next three sections.

6. Euclidean invariants

To construct a Euclidean invariant in the frequency domain, we need the following theorem, whose proof is straightforward.

**Theorem 9** (Rotational symmetry of the Taylor invariant). The Taylor invariant \( F_c \) is rotationally symmetric.

Therefore, we have the following Euclidean invariant.

**Definition 22** (Taylor–Mellin–RT invariant). The Taylor–Mellin–RT invariant is
\[ M_{RT} = M_R \circ F_c. \]

By Theorems 6 and 9, \( M_{RT} \) is a complete invariant with respect to \( RT \).

The other translation invariant in the frequency domain is the Hessian invariant \( F_H \). \( F_H \) is, however, not rotationally symmetric. In order to have a rotationally symmetric translation invariant, we need to modify the Hessian invariant as follows.

**Definition 23** (Symmetric Hessian invariant). The symmetric Hessian invariant is given by
\[ H(\omega_x, \omega_y) = \Omega \times F_H(\omega_x, \omega_y), \]
where

$$
\Omega = \begin{bmatrix}
\omega_x^2 + \omega_y^2 & 0 & 0 & 0 \\
0 & \omega_x^2 & 2\omega_x\omega_y & \omega_y^2 \\
0 & \omega_x\omega_y & \omega_y^2 - \omega_x^2 & -\omega_x\omega_y \\
0 & \omega_y^2 & -2\omega_x\omega_y & \omega_x^2 \\
\end{bmatrix}.
$$

The following theorem states that $H$ is a complete invariant.

**Theorem 10** (Complete invariant, $H$). The symmetric Hessian invariant $H$ is a complete invariant with respect to $T$.

**Proof.** $H$ is invariant with respect to $T$, because for all $T_{\alpha,\beta} \in T$, by the fact that $F_H$ is invariant with respect to $T$,

$$H(\omega_x, \omega_y) \circ T_{\alpha,\beta} = \Omega \times F_H(\omega_x, \omega_y) \circ T_{\alpha,\beta} = \Omega \times F_H(\omega_x, \omega_y) = H(\omega_x, \omega_y).$$

To show the completeness, note that $\Omega$ is invertible except at $(\omega_x, \omega_y) = (0, 0)$. Thus, $F_H$ can be recovered from $H$ except at $(\omega_x, \omega_y) = (0, 0)$. Since $F$ (and hence $F_H$) is entire, due to the assumption of the bounded support and finite energy for the original image $f$, $F_H(0,0)[f]$ can be determined from its limit. Hence the completeness of $F_H$ implies the completeness of $H$. □

**Theorem 11** (Rotational symmetry of the symmetric Hessian invariant). The symmetric Hessian invariant $H$ is rotationally symmetric.

**Proof.** By the properties of Fourier representation,

$$A(\omega_x, \omega_y) \circ R_\theta = R_\theta(\omega_x, \omega_y) \circ A, \quad \Phi(\omega_x, \omega_y) \circ R_\theta = R_\theta(\omega_x, \omega_y) \circ \Phi.$$

Therefore, it can be calculated that

$$F_H(\omega_x, \omega_y) \circ R_\theta = \Theta \times F_H(\omega_x, \omega_y),$$

where

$$\Theta = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \theta^2 & -2 \sin \theta \cos \theta & \sin \theta^2 \\
0 & \sin \theta \cos \theta & \cos \theta^2 - \sin \theta^2 & -\sin \theta \cos \theta \\
0 & \sin \theta^2 & 2 \sin \theta \cos \theta & \cos \theta^2 \\
\end{bmatrix}.$$

It can then be shown that $\Omega \times \Theta = R_\theta \circ \Omega$. Hence,

$$H(\omega_x, \omega_y) \circ R_\theta = \Omega \times (F_H(\omega_x, \omega_y) \circ R_\theta) = \Omega \times \Theta \times F_H(\omega_x, \omega_y) = (R_\theta(\omega_x, \omega_y) \circ \Omega) \times F_H(\omega_x, \omega_y) = R_\theta(\omega_x, \omega_y) \circ H. \quad \square$$

Hence the other Euclidean invariant can be defined as follows.

**Definition 24** (Hessian–Mellin-RT invariant). The Hessian–Mellin-RT invariant is given by

$$H_{RT} = M_R \circ H.$$
By Theorems 6, 10 and 11, \( M_{RT} \) is a complete invariant with respect to \( RT \).

7. Translation and scaling invariants

To construct translation and scaling invariants in the frequency domain, we need the property of reciprocal scaling. The Taylor invariant \( F_t \), however, does not satisfy this property. So we modify the Taylor invariant to the following Laplacian invariant.

**Definition 25** (Laplacian invariant). Let \( \omega_x, \omega_y \) = \( \omega_x^2 + \omega_y^2 \). The Laplacian invariant is given by

\[
L(\omega_x, \omega_y) = r(\omega_x, \omega_y) \times F_c(\omega_x, \omega_y).
\]

The name Laplacian invariant comes from the fact that it can be derived from the Fourier transform of the Laplacian of \( f(x, y) \). \( L \) is a complete invariant as shown in the following theorem, whose proof is similar to that for \( H \).

**Theorem 12** (Complete invariant, \( L \)). The Laplacian invariant \( L \) is a complete invariant with respect to \( T \).

**Theorem 13** (Reciprocal scaling of the Laplacian invariant). The Laplacian invariant \( L \) is reciprocally scaled.

**Proof.** By the definitions and previous results,

\[
L \circ S_p = r \times F_c \circ S_p = r \times \frac{1}{\rho^2} \times S_{1/\rho} \circ F_c = S_{1/\rho} \circ r \times F_c = S_{1/\rho} \circ L. \tag*{\QED}
\]

Now, we can define a frequency domain translation and scaling invariant.

**Definition 26** (Laplacian–Mellin-ST invariant). The Laplacian–Mellin-ST invariant is given by

\[
M_{ST} = M_S \circ L.
\]

By Theorems 7, 12 and 13, \( M_{ST} \) is a complete invariant with respect to \( ST \).

Another way to construct translation and scaling invariants is to combine the Mellin-S invariant with the symmetric Hessian invariant. To do this, we first need the following theorem.

**Theorem 14** (Reciprocal scaling of the symmetric Hessian invariant). The symmetric Hessian invariant \( H \) is reciprocally scaled.

**Proof.** It is similar to the proof of Theorem 13. \( \tag*{\QED} \)

Therefore, we have the following translation and scaling invariant.

**Definition 27** (Hessian–Mellin-ST invariant). The Hessian–Mellin-ST invariant is given by

\[
H_{ST} = M_S \circ H.
\]

By Theorems 7, 10 and 14, \( H_{ST} \) is a complete invariant with respect to \( ST \).
8. Similarity invariants

We now discuss similarity invariants. By Theorem 13, the Laplacian invariant is reciprocally scaled. It is also rotationally symmetric as shown in the following theorem.

**Theorem 15** (Rotational symmetry of the Laplacian invariant). The Laplacian invariant $L$ is rotationally symmetric.

**Proof.** It follows from the fact that 
\[ R_\theta \circ (\omega_x^2 + \omega_y^2) = \omega_x^2 + \omega_y^2. \]
\[ \square \]

Hence, we can construct the following similarity invariant.

**Definition 28** (Laplacian–Mellin-RST invariant). The Laplacian–Mellin-RST invariant is given by 
\[ M_{RST} = M_{RS} \circ L. \]

By Theorems 8, 13 and 15, $M_{RST}$ is a complete invariant with respect to RST.

On the other hand, the symmetric Hessian invariant is rotationally symmetric and reciprocally scaled by Theorems 11 and 14. Therefore we have the following similarity invariant.

**Definition 29** (Hessian–Mellin-RST invariant). The Hessian–Mellin-RST invariant is given by 
\[ H_{RST} = M_{RS} \circ H. \]

By Theorems 8, 11 and 14, $H_{RST}$ is a complete invariant with respect to RST.

9. Conclusion

We first reviewed the Taylor and Hessian invariants proposed in (Lin, 1993):

*Translation invariants:* \( F_C, \quad F_H. \)

Then we presented the remaining basic invariants. They are summarized as follows:

*Scaling invariants:* \( M_S. \)

*Rotation invariants:* \( M_R. \)

*Rotation-scaling invariants:* \( M_{RS}. \)

All these invariants are complete invariants. We then modified the translation invariants to obtain the following invariants which satisfy the properties of rotational symmetry and reciprocal scaling:

*Translation invariants:* \( L, \quad H. \)

Based on the new invariants, we can construct the hybrid invariants. They are summarized as follows:

*Euclidean invariant:* \[ M_{RT} = M_R \circ F_C, \quad H_{RT} = M_R \circ H, \]

*Translation and scaling invariant:* \[ M_{ST} = M_S \circ L, \quad H_{ST} = M_S \circ H, \]

*Similarity invariant:* \[ M_{RST} = M_{RS} \circ L, \quad H_{RST} = M_{RS} \circ H. \]
We can reconstruct the complete images, if needed, modulo translation, rotation and scaling from these complete invariants in a manner similar to that discussed in Section 3.

Acknowledgements

This research is supported in part by the National Science Foundation under grant ECS-9315344. We would like to thank Andrew Harris for his technical assistance.

References