Comput. Math. Applic. Vol. 13, No. 4, pp. 363–365, 1987
Printed in Great Britain.

0097-4943/87 $3.00 + 0.00

AN APPLICATION OF GROUP REPRESENTATION THEORY TO PICTURE RECOGNITION

C. ASMUTH
Department of Mathematics and Computer Science, Rutgers University-Newark, Newark, NJ 07102, U.S.A.

(*Received* 3 *May* 1986)

Communicated by E. Y. Rodin

Abstract—An algorithm is given for producing numerical values associated to a picture (assumed to be in a neutral background) which are independent of rigid transformation of the picture. The computations are developed in a very general way based on unitary representation of groups of geometric transformations.

1. INTRODUCTION

This paper proposes a technique for machine recognition of a single image. This technique will work independently of the position and orientation of the image. The image is considered to be a function on the plane $\mathbb{R}^2$. From this function we compute some numerical invariants, i.e. constants which are unaffected by rigid transformations of the image. These values will (most of the time) be different for different images. The application of this method would be where the recognition algorithm would encounter members of a finite set of images. For each image in the set, the invariant values would be computed ahead of time. The recognition process would consist of first computing the invariants for the current image and then matching those values with one of the sets of pre-computed values.

The theory behind the technique, as outlined in the next sections, is that of group representations. The group in question is naturally that of all rigid transformations of the plane. Such theory is closely related to concepts of Fourier series and Fourier transforms. These, in fact deal with representation theory of the complex unit circle multiplicative group and the additive group of the real line, respectively. What makes things more complicated here is that, unlike these two groups, the operation in the group of rigid transformations is not commutative. (Consider the process of first rotating by $\pi/4$ centered at the origin, then translating one unit in the positive x-axis direction. Compare the results of this to the results of the operations in reverse order.)

We now give a formula for computing the numerical invariants. Let f be a positive real-valued function of two real variables which represents the image. Then for each $r \geqslant 0$ we obtain the invariant value

$$F(r) = \int_0^{2\pi} \left\{ \iint_{\mathbb{R}^2} f(x, y) \exp[iR(x \cos\theta + y \sin\theta)] \, dx \, dy \right\}^2 d\theta.$$

What follows is a development of the theory behind this formula.

SECTION 2

Here we develop the theoretical foundation from which the recognition technique is derived. Our discussion is quite general and applies to image recognition problems in any dimension so long as the specific calculations are feasible.

To start with, let X be a measure space and let G be a topological group with right Haar measure (right translation invariant) dg. We let G act on X; the action of $g \in G$ sends an element $x \in X$ to xg. We let x_0 be some fixed reference point of X and require the following:

(1) For any $x \in X$, there is some $g \in G$ such that $x = x_0 g$ (G acts transitively).

(2) If it is a complex square integrable function [$f \in L^2(x)$], there is a unique $\tilde{f}$ in $L^2(G)$ such that $f(g) = f(x_0 g)$.

Of course, the example we are interested in is where X is the plane $\mathbb{R}^2$ and G is some group of rigid transformations of the plane.

For completeness we now include some elementary definitions and results.

Definition 2.1. A unitary matrix representation U of degree n of the group G is, for each g in G, a choice of an $n \times n$ unitary matrix $U(g)$ such that for any pair of elements g_1 and g_2 in G we have $U(g_1)\,U(g_2) = U(g_1 g_2)$.

Definition 2.2. Two unitary representations U_1 and U_2 of the same degree are said to be equivalent if there is some unitary matrix P such that for all g in G, $P^{-1}U_1(g)P = U_2(g)$.

Definition 2.3. A unitary matrix representation U of degree n is said to be irreducible if there is no non-trivial subspace $V \subseteq \mathbb{C}^n$ such that for any $g \in G$, $U(g)V \subseteq V$.

We let $\hat{G}$ denote a complete set of irreducible inequivalent unitary matrix representations of G. For the present, we will only consider groups where all unitary representations have finite degree. For each U in $\hat{G}$, let $U_{ij}(g)$ be the ij-entry of the matrix $U(g)$. For $\tilde{f}$ in $L^2(G) \cap L^1(G)$ we set

$$L^U_{ij}(\tilde{f}) = \int_G \tilde{f}(g)\,\overline{U_{ij}(g)}\,\mathrm{d}g.$$

Representation theory tells us that such a function $\tilde{f}$ is entirely determined by the numbers L^U_{ij}. A simple example to keep in mind is when G is the multiplicative group of the complex unit circle. Here $\hat{G} = \{T_n: n \in \mathbb{Z}\}$; each $T_n(e^{i\theta}) = e^{in\theta}$. All T_ns are degree one; the number $L^{T_n}(\tilde{f})$ is just the nth Fourier coefficient of $\tilde{f}$.

We return now to the picture problem. A picture will be considered to be a positive real-valued function f on X. For $h \in G$, set $f^h(x) = f(xh)$. Thus the picture represented by f^h is the same as that of f only moved according to the action of the element h in G. The key idea is to follow the effect of h through to the computation of each $L^U_{ij}[(\tilde{f^h})]$. Since $f(g) = \tilde{f}(x_0 g)$, $(\tilde{f^h})(g) = f^h(x_0 g) = f(x_0 gh) = \tilde{f}(gh)$. If the degree of U is n we have

$$\begin{aligned} L^U_{ij}[(\tilde{f^h})] &= \int_G (\tilde{f^h})(g)\,\overline{U_{ij}(g)}\,\mathrm{d}g \\ &= \int_G \tilde{f}(gh)\,\overline{U_{ij}(g)}\,\mathrm{d}g \\ &= \int_G \tilde{f}(g)\,\overline{U_{ij}(gh^{-1})} \\ &= \sum_{k=1}^{n} \int_G \tilde{f}(g)\,\overline{U_{ik}(g)\,U_{kj}(h^{-1})}\,\mathrm{d}g \\ &= \sum_{k=1}^{n} L^U_{ik}(\tilde{f})\,\overline{U_{kj}(h^{-1})}. \end{aligned}$$

Now let $\mathscr{L}^U(f)$ be the $n \times n$ matrix whose entries are the $L^U_{ij}(\tilde{f})$s. The computation above shows that $\mathscr{L}^U(f^h) = \mathscr{L}^U(f)\,\overline{U(h^{-1})}$. For v and w in $\mathbb{C}^n$, let $\langle v, w\rangle$ denote the standard Hermitian form. Assume that v and we are any two rows of $\mathscr{L}^U(f)$. Let v^h and w^h be rows of $\mathscr{L}^U(f^h)$. Then since $\overline{U(h^{-1})}$ is unitary, we have $\langle v, w\rangle = \langle v^h, w^h\rangle$. These are position invariant values for the image f.

SECTION 3

We now develop the specific case of images in the plane. Let $X = \mathbb{R}^2$ and let G be the group of rigid transformations consisting of all translations and all rotations centered at the origin through angles which are multiples of $2\pi/N$ for some fixed N. Let $\mathscr{R} = \{r_0, \ldots, r_{N-1}\}$ be these rotations where r_k is a rotation through the angle $2\pi K/N$. $\mathscr{T}$ will denote the subgroup of all translations; its elements will be given as vectors in $\mathbb{R}^2$. If g is an element of G, it may be uniquely written $g = r_k w$ where w is a member of $\mathscr{T}$. The point (0, 0) will serve as x_0.

From elementary group representation theory we know that there are two kinds of elements in $\hat{G}$. They are:

(1) Representations of degree one. These will be trivial on $\mathscr{T}$.
(2) Representations of degree N. These are of the form U^w where $w \in \mathbb{R}^2$. U^w and U^v are equivalent when w and v are conjugate by $\mathscr{R}$.

It turns out that if U is of the first sort, then all $\mathscr{L}^U(f)$s are zero for image f. Therefore we consider only the second sort. For w in $\mathbb{R}^2$, let w_k be the vector w moved through the rotation $2\pi k/N$. For $g = r_k v$ in G we set

$$U^w(g) = \left(\begin{array}{c|c} 0 & I_{N-k} \\ \hline I_k & 0 \end{array}\right)\begin{pmatrix} d_0 & & 0 \\ & \ddots & \\ 0 & & d_{N-1} \end{pmatrix}.$$

Here I_s denotes the $s \times s$ identity matrix and $d_m = \exp(iw_m \cdot v)$

Now let f be an image function whose corresponding function on G is $\tilde{f}$. Then:

$$L_{ij}^{U^w}(\tilde{f}) = \sum_{k=0}^{N-1} \int_{v \in \mathbb{R}^2} \tilde{f}(r_{kv})\, \overline{U_{ij}^w(r_k v)}\, \mathrm{d}r$$

$$= \sum_{k=0}^{N-1} \int_{v \in \mathbb{R}^2} \tilde{f}(v) \sum_{m=1}^{N} \overline{U_{im}^w(r)\, U_{mj}(v)}\, \mathrm{d}r \quad (\text{since } x_0 r_k = x_0)$$

$$= N \int_{v \in \mathbb{R}^2} \tilde{f}(v)\, U_{ij}^w(v)\, \mathrm{d}v.$$

This last formula is just a multiple of the usual 2-D Fourier transform $(\tilde{f})\hat{}$ of $\tilde{f}$ at w_j. Thus for each $w \in \mathbb{R}^2$ we compute an invariant of the form

$$F(w) = \sum_{j=0}^{N-1} \left| (\tilde{f})\hat{}(w_j) \right|^2.$$

The formula in Section 1 is obtained by letting $N \to \infty$ and using only ws of the form (R, O) to parametrize the representation of $\hat{G}$.

REFERENCE

1. L. Dornhoff, *Group Representation Theory*, Part A. Dekker, New York (1971).