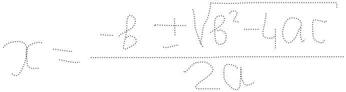
Orientation-Independent Recognition of Handwritten Characters with Integral Invariants

Oleg Golubitsky, Vadim Mazalov and Stephen M. Watt

Ontario Research Center for Computer Algebra
Department of Computer Science
University of Western Ontario

Introduction

- Our objective is to recognize handwritten mathematics in pen-based environment.



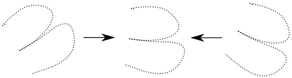
$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

- What we have: 97.5% recognition rate for standalone mathematical symbols achieved with approximation of coordinate functions with Legendre-Sobolev polynomials.

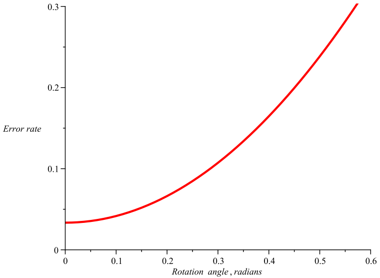
- What we want: extend this result to rotated and, possibly, sheared symbols.

Rotation

- Commonly occurs in practice.



- Decreases recognition rate in our previous algorithm with quadratic dependence on the rotation angle.



Challenge

- Question

Is it possible to make an algorithm independent of rotation?

- Answer

Yes. Describe the curve in terms of rotation invariant functions:

- Integral invariants [1]
- Geometric moments [2]

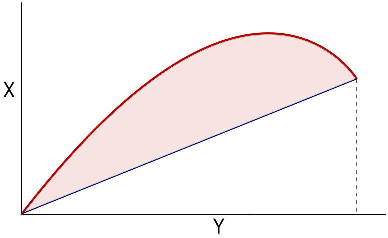
Integral Invariants

- Provide an elegant approach to planar and spatial curve classification under affine transformations.
- Relatively insensitive to small perturbations, as opposed to differential invariants.
- Can be computed online, i.e. when the curve is being written, with minor overhead after pen-up.

As the name suggests, integral invariants are constructed from quantities obtained by integration.

Geometric Representation

- Invariant $I_0(\lambda)$ is the radius to a point on a curve.
- Invariant $I_1(\lambda)$ is the area between the curve and a secant.



Functional Representation

- I_0 and I_1 are invariant under rotation and may be given in terms of coordinate functions $X(\lambda)$ and $Y(\lambda)$:

$$I_0(\lambda) = \sqrt{X^2(\lambda) + Y^2(\lambda)} = R(\lambda),$$
$$I_1(\lambda) = \int_0^{\lambda} X(\tau) dY(\tau) - \frac{1}{2} X(\lambda) Y(\lambda)$$

  where $X(\lambda)$, $Y(\lambda)$ are parameterized by Euclidean arc length.

- Note: $I_0(\lambda)$ is invariant under the action of the special orthogonal group $SO(2)$, while $I_1(\lambda)$ is invariant under $SL(2)$, the group of special linear transformations.

Functional Approximation

- We represent the functions I_0 and I_1 by orthogonal series coefficients.
- An inner product $\langle f, g\rangle$ gives the orthogonal basis $\{\phi_0(\lambda), ..., \phi_d(\lambda)\}$ on a subspace of polynomials by GS orthogonalization of $\{\lambda^i\}$.
- Can approximate

$$f(\lambda) \approx \sum_{i=0}^{d} \alpha_i \phi_i(\lambda) \qquad \alpha_i = \langle f, \phi_i\rangle$$

- We use a Legendre-Sobolev i.p. so we can measure distance in the first jet space.

$$\langle f, g\rangle = \int_{-1}^{1} f(\lambda)g(\lambda)\mathrm{d}\lambda + \mu \int_{-1}^{1} f'(\lambda)g'(\lambda)\mathrm{d}\lambda$$

Numerical Representation

- After approximation of coordinate functions with truncated Legendre-Sobolev polynomial series, invariants $I_0(\lambda)$ and $I_1(\lambda)$ take form

$$I_0(\lambda) = \sqrt{\left(\sum_{i=1}^{d} \bar{x}_i P_i(\lambda)\right)^2 + \left(\sum_{i=1}^{d} \bar{y}_i P_i(\lambda)\right)^2},$$

$$I_1(\lambda) = \sum_{i,j=1}^{d} \bar{x}_i \bar{y}_j \left[\int_0^{\lambda} P_i(\tau) P_j'(\tau) d\tau - \frac{1}{2} P_i(\lambda) P_j(\lambda)\right]$$

where P_i is the i-th Legendre-Sobolev polynomial.

Remark about Shear

Can be treated with integral invariants similar to rotation, but there are associated difficulties:

- Selection of the appropriate shear-invariant parameterization of coordinate functions.
- Requirements of careful analysis of transformation limits to avoid blending classes.

Figure: Ambiguity, introduced by shear and rotation

In practice shear is seen most on tall, thin symbols and this can to an extent be corrected by rotation.

Geometric Moments

- Of special interest for the purpose of online curve classification under pressure of computational constraints, since they are easy to calculate, while invariant under scaling, translation and rotation.

- A $(p+q)$-th order moment of f can be expressed as

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y)$$

- Since coordinates in our algorithm are already normalized with respect to size and rotation, we work with moments directly.

Numerical Representation

- The $(p+q)$-th moment of a sample's coordinates can be expressed as

$$m_{pq}(\lambda_\ell) = \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} X^p(\lambda_i) Y^q(\lambda_j) f\big(X(\lambda_i), Y(\lambda_j)\big)$$

where $X(\lambda_i)$ and $Y(\lambda_i)$ are coordinates X and Y at sample point i; $f\big(X(\lambda_i), Y(\lambda_j)\big) = \sqrt{X^2(\lambda_i) + Y^2(\lambda_j)}$.

- Moment invariants have form

$$\begin{aligned} M_0(\lambda) &= m_{00}(\lambda), \\ M_1(\lambda) &= m_{20}(\lambda) + m_{02}(\lambda), \\ M_2(\lambda) &= (m_{20}(\lambda) - m_{02}(\lambda))^2 + 4m_{11}^2(\lambda). \end{aligned}$$

Pre-classification

- Coordinate functions $X(\lambda)$ and $Y(\lambda)$, parameterized by arc length, are approximated with Legendre-Sobolev polynomials. These polynomials are orthogonal with respect to the inner product

$$\langle f, g \rangle = \int_a^b f(\lambda) g(\lambda) d\lambda + \mu \int_a^b f'(\lambda) g'(\lambda) d\lambda.$$

- Having a coefficients vector $(x_0, \ldots, x_d, y_0, \ldots, y_d)$ of the approximation for $X(\lambda)$ and $Y(\lambda)$, the sample can be normalized with respect to position and size by omitting the first point (x_0, y_0) and normalizing the vector $(x_1, \ldots, x_d, y_1, \ldots, y_d)$.

Classification

The following algorithms are introduced:

- Classification with integral invariants (CII)
- Classification with coordinate functions and integral invariants (CCFII)
- Classification with coordinate functions and moment invariants (CCFMI)

Classification with integral invariants (CII)

- Integral invariants are considered as the curves to be approximated, yielding vectors

$$(\bar{I}_{0,1}, \ldots, \bar{I}_{0,d}, \bar{I}_{1,1}, \ldots, \bar{I}_{1,d}).$$

- Classification is based on evaluation of the distance from a sample to convex hulls of the nearest neighbours and selection of the class with the smallest distance.
- The algorithm does not depend on the number of classes, since only one class is considered.

Classification with coordinate functions and integral invariants (CCFII)

- Integral invariants are considered as the curves to be approximated, just like in CII.
- As an alternative to CII, coefficients of approximation of integral invariant functions are used to select the closest N candidate classes.
- N can be chosen empirically to ensure high probability of the correct class been selected.

Classification with coordinate functions and integral invariants (CCFII), cont.

- The correct class is determined as the solution to the minimization problem of the distance to corresponding classes with respect to rotation

$$\min_{\alpha}\left(\sum_k (X_k - (x_k \cos\alpha + y_k \sin\alpha))^2 + \sum_k (Y_k - (-x_k \sin\alpha + y_k \cos\alpha))^2\right)$$

- The stationary point has the form:

$$\alpha = \arctan\left(\frac{\sum_k (X_k y_k - Y_k x_k)}{\sum_k (X_k x_k + Y_k y_k)}\right)$$

Classification with coordinate functions and moment invariants (CCFMI)

- Moment invariants are considered as the curves to be approximated, similar to CII.
- Coefficients of approximation of moment invariant functions are used to select the closest N candidate classes.
- The right class is determined as the solution to minimization problem, just like in CCFII.

Experimental Setting

- Our dataset comprised 50,703 handwritten mathematical symbols from 242 classes.
- Testing is implemented in 10-fold cross-validation (the dataset was randomly divided into 10 parts, preserving proportions of class sizes).
- Normalized Legendre-Sobolev coefficient vectors of coordinate functions of randomly rotated symbols, as well as coefficients of integral invariants, were pre-computed for all symbols.

CCFII

Table: Presence of the correct class within top N classes, %

$N=1$	2	3	4	5	6	7	10	15	**20**	25
87.9	95.1	96.8	97.7	98.3	98.7	98.9	99.4	99.5	**99.5**	99.5

Table: Error rate, depending on the number of nearest neighbours, %

angle (radians)	$K=8$	10	12	14	16	18	19	**20**	21	22
0	4.4	3.9	4.2	4.0	3.9	3.9	3.8	**3.7**	3.8	3.8
0.3	6.2	5.7	5.7	5.4	5.4	5.4	5.3	**5.3**	5.4	5.4
0.5	7.4	6.9	6.8	6.7	6.6	6.5	6.4	**6.4**	6.5	6.5
0.7	8.5	7.9	7.7	7.6	7.4	7.4	7.2	**7.2**	7.3	7.4
0.9	9.3	8.8	8.6	8.3	8.2	8.2	8.2	**8.1**	8.2	8.2
1.1	9.6	9.0	8.7	8.6	8.4	8.4	8.2	**8.2**	8.4	8.4
average	7.5	7.0	7.0	6.8	6.6	6.6	6.5	**6.5**	6.6	6.6

CCFMI

Table: Presence of the correct class within top N classes, %

$N = 1$	2	3	4	5	10	20	30	40	**50**	55
51.5	68.3	77.2	82.2	85.9	95.3	98.8	98.9	99.0	**99.0**	99.0

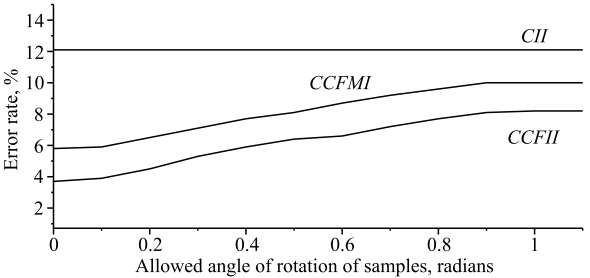
Table: Error rate, depending on the number of nearest neighbours, %

angle (radians)	$K = 8$	10	12	14	16	18	20	21	**22**	23
0	7.0	6.6	6.4	6.2	6.1	6.1	5.9	5.8	**5.8**	6.0
0.3	8.0	7.8	7.6	7.4	7.2	7.1	7.0	7.2	**7.1**	7.2
0.5	9.3	9.1	8.9	8.5	8.3	8.3	8.2	8.2	**8.1**	8.3
0.7	10.4	10.1	9.9	9.5	9.4	9.2	9.1	9.2	**9.2**	9.3
0.9	11.5	11.1	10.8	10.4	10.2	10.2	10.1	10.0	**10.0**	10.0
1.1	11.4	11.1	10.7	10.4	10.2	10.1	10.1	10.0	**10.0**	10.0
average	9.6	9.3	9.1	8.7	8.6	8.5	8.4	8.4	**8.4**	8.5

Comparison of results

Table: Error rates of CII, CCFII and CCFMI

α, rad.	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	1.0	1.1
CII	12	12	12	12	12	12	12	12	12	12
CCFII	**3.7**	**3.9**	**4.5**	**5.3**	**5.9**	**6.4**	**6.6**	**7.2**	**8.2**	**8.2**
CCFMI	5.8	5.9	6.5	7.1	7.7	8.1	8.7	9.2	10	10



Analysis of Results

- CCFII performs noticeably better, while requiring less computation.
- As expected, we observed an increase in error rate with the rotation angle for CCFII and CCFMI. Typical misclassifications arise when distinct symbols have similar shape and are normally distinguished by their orientation, for example “1” and “/”, “+” and “$\times$”, “U” and “$\subset$”.
- This result can still be improved, but in a different setting.

Conclusion

- Recognition rate can be improved, if a system could consider the tendency to write characters in similar orientations and restrict the range of angles for nearby symbols.

- For now, our results confirm that integral invariants are a suitable instrument in the recognition of handwritten characters, when orientation is uncertain, and provide an acceptable recognition rate.

References

1. S. Feng, I. Kogan and H. Krim, Classification of Curves in 2D and 3D via Affine Integral Signatures, to appear in *Acta Appl. Math.*, 2008, `http://arxiv.org/abs/0806.1984`.
2. M.K. Hu, Visual Pattern Recognition by Moment Invariants, IRE Transactions on Information Theory, Vol. IT-8, pp. 179-187, 1962.