1H－5

An Experimental System for the Recognition of Handwritten Uygur Characters

Anwer.YMIN　Toshio KAWASHIMA　Yoshinao AOKI
Faculty of Engineering, Hokkaido University

1. Introduction

An exprimental on-line hand written Uygur character recognition system is described. In the system,input patterns are translated into "return pattern", a normalized description of input, preserving topological characteristics. Features that are specific to each alphabet, called principal features,are extracted from a series of translated main strokes which forms a continuous return pattern. Second strokes, a set of secondary segments,further divide the character class by their position and the number of the strokes . The identified characters are checked according to the connection rule of Uygur writing. The recognition experiment of our Uygur recognition system is 94% correct.

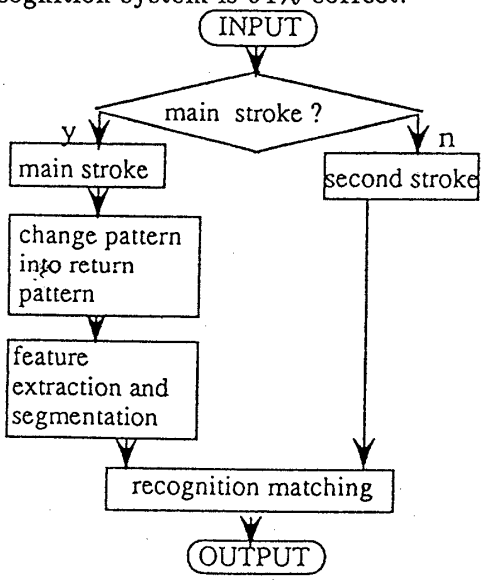

Fig.1 System of uygur character recognition.

手書きウイグル文字認識システム
アニワル・イミン、川嶋　稔夫、青木　由直
北海道大学工学部
札幌市北区北１３条西８丁目

2. Preprocessing

The preprocessing of uygur character recognition in the system consists of two parts ;speed filter and direction filter.

Speed filter We developed a trim filter for uygur character recognition. It relies on pen dynamics,namely velocity and acceleration. The mass of a human hand holding of the pen and the force that can be applied to it limits acceleration. Large acceleration is from large velocity change or angle changes. This provides a convenient mean to remove the unnecessary points.

Direction filter The second-level filter is based on the direction of a stroke . This filter limits the direction of input pattern, because we use only four directions to describe the characters. This filter preserves the most of topological characteristics while it removes redundancy from inputs. Examples of input pattern and their return patterns are shown in fig.2.

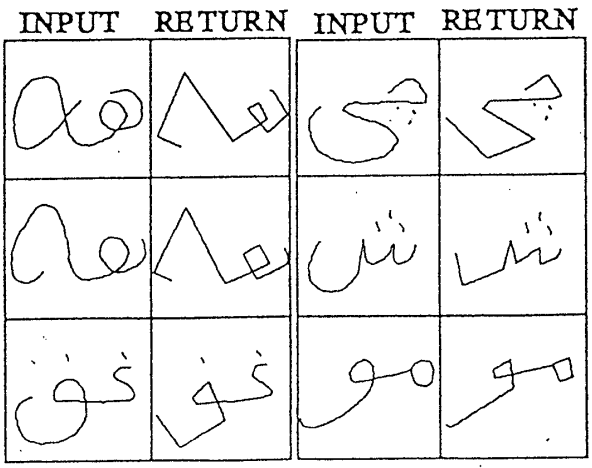

Fig.2 Input/Return patterns

3. Classification of characters and

representation of principal features

We use two types of features to cassify transformed patterns. The first is the type of a main stroke which embodys the skelton of a character,and the second is the position and number of second strokes. In the first step, the system extracts the principal features of main strokes from a series of main strokes as shown in fig.3, and then,classify the features into several classes from their topological structure as shown in fig.4. The process corresponds to the left flow in fig.1. If the character has second strokes, the system identifies the character class from the number and type of second strokes and from their relative position to the main stroke. The types of second strokes are short segment and dots.

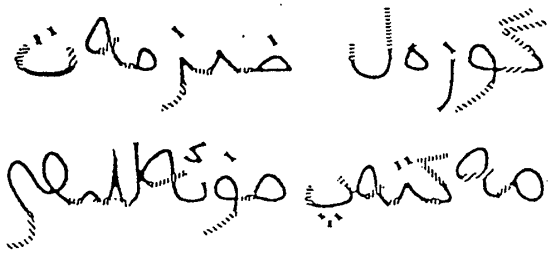

Fig.3 Principal features of characters(solid line)

FEATURE	INPUT PATTERN	RETURN PATTERN	INPUT PATTERN	RETURN PATTERN
LOOP				
CUSP				
HUMP				
CLOSURE				

Fig.4 Typical examples of principal features

4. The hierarchical recognition of uygur character

The whole recognition flow is shown in fig.4. The flow is divided into four stages.

Pre-pattern stage extracts principal features from a series of main strokes written in cursive style.This stage outputs the class of main stroke. Second-pattern stage extracts second strokes and reports the class and number of the second strokes. Block-pattern integrates P-pattern and S-pattern to identify an Uygur alphabet. The correspondence between P and S-patterns are determined by their relative position.Uygur-pattern stage checks the connection between main strokes recognized by block-pattern stage from the rule of cursive writing in Uygur.

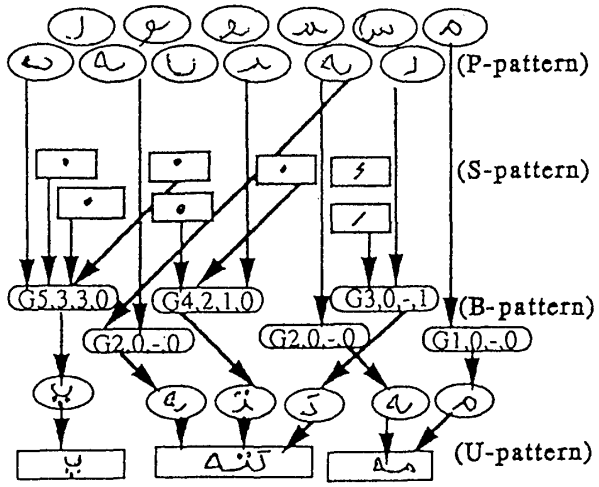

Fig.4 Hierarchical uygur character recognition.

5. Conclusion

The outline of our Uygur recognition system has described. The cursive script of Uygur can be analyzed by four step analysis. The system,however, fails when the input is extremely deformed, because principal features are fused into a smooth line segment. Our future study is concentrated on the deformation rule of Uygur characters.

Reference

R.Plamondon and C.G Leedham,editors. Computer Processing of Handwriting.Word Scientific,1990.