

HAND-DRAWN OBJECTS RECOGNITION

Jakub Křístek

Master Degree Programme (2), FIT BUT nebo FEEC BUT

E-mail: xkrist05@stud.feec.vutbr.cz

Supervised by: Oto Janoušek

E-mail: janouseko@feec.vutbr.cz

Abstract: This work deals with recognition of hand-drawn objects traced by children with mental disorders. The aim is to classify object's geometrical primitives into classes and display them along with the idealized shape of the input object. Level of mental retardation is determined by the variance of the input (drawn) object from idealized shape of the object (artwork).

Keywords: Object recognition, hand-drawn objects, feature description methods

1. ÚVOD

Způsob kresby, resp. míru odchylky nakresleného objektu od předlohy, lze využít k diagnostice psychických poruch osob. Kresebné testy jsou využívány zejména v diagnostice dětí, neboť se předpokládá, že děti kreslit baví a není třeba je podrobovat složitým psychologickým testům. Kresebné metody eliminují problém porozumění textu, v případě kreslení dle předlohy každé dítě pochopí zadaný úkol.

Diagnostická výtěžnost kresebných testů je však limitována subjektivním hodnocením terapeuta, který provádí vyhodnocení kresby. Kromě toho je administrace kresebného testu a jeho hodnocení časově velmi náročné. Automatické rozpoznání kresby a míry její distorze obě omezení překonává.

Postupy popisované v této práci předpokládají, že bylo provedeno předzpracování kresby. Předzpracování spočívá ve vektorizaci čar, ze kterých se kresba skládá, a v určení pořadí jejich vzniku. Každý objekt v kreslené scéně je po předzpracování tvořen uceleným souborem vektorů, které jsou charakterizovány počátečním bodem, směrovostí a příslušnou délkou. Soubory vektorů jsou následně popsány svými příznaky, na základě jejichž analýzy je objekt klasifikován do stanovených tříd objektů.

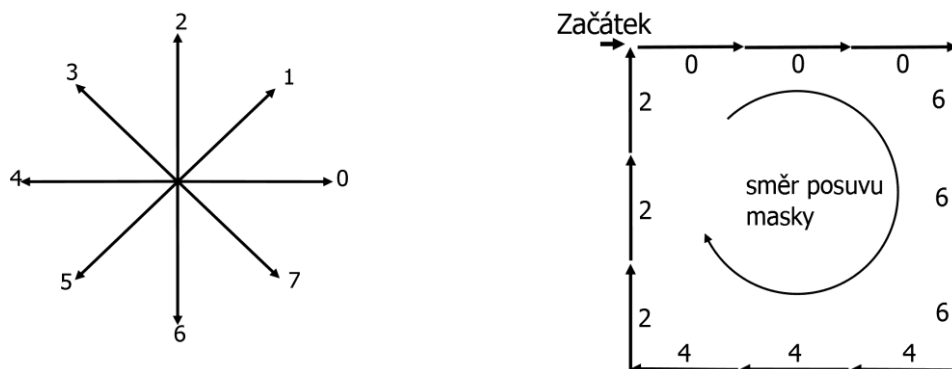
Rozpoznávání ruční kresby v této práci je omezeno pouze na základní geometrické obrazce, tj. čtverec, obdélník, trojúhelník a kružnice, ze kterých pak spojováním vznikají obrazce složitější. Na těchto základních tvarech lze otestovat navrženou metodu a tu pak aplikovat na celou sadu objektů.

2. METODA POPISU A KLASIFIKACE

Popsané objekty jsou reprezentovány vektorem příznaků, který má charakter číselného řetězce, a slouží k následné klasifikaci objektů. Ideálně je hledán takový popis, který je invariantní vůči rotaci, posunutí i velikosti objektů. [1]

Protože se jedná o popis samostatných objektů, nikoliv rozpoznávání primitiv v obraze, spadá zvolená metoda do oblasti příznakových metod. Pro popis lze použít metodu Freemanova vrcholového nebo pixelového kódu, Fourierovy deskriptory, deskriptory B-spline křivky nebo momentovou metodu. Protože je třeba měnit velikost idealizovaného objektu v závislosti na velikosti vstupního (nakresleného) objektu a také je pro následnou klasifikaci vyžadována rotační i translační invariance, byla zvolena metoda Freemanova řetězového pixelového kódu (dále jen Freemanův kód).

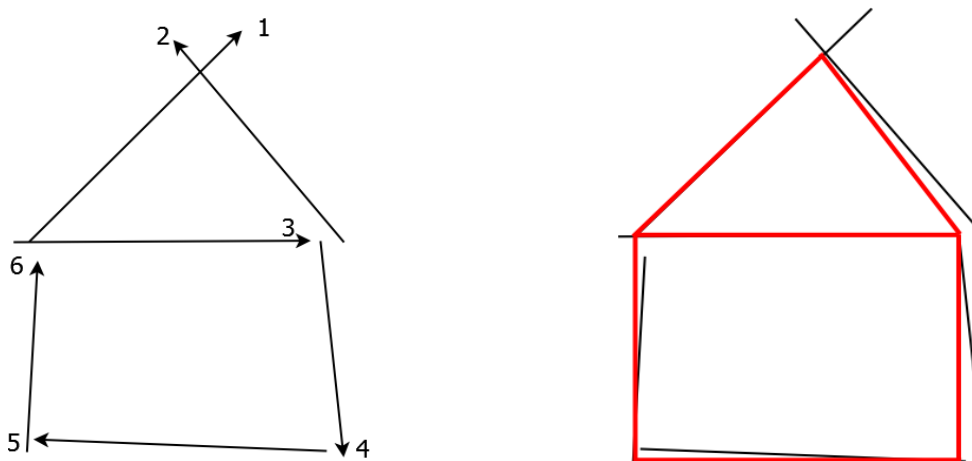
Princip popisu objektu pomocí Freemanova kódu spočívá v určení hodnot příznaků z hraniční reprezentace objektu, která je v případě kreslených objektů jednoznačně vymezena. Každému hraničnímu pixelu je přiřazeno příslušné číslo ze schématu (**Obrázek 1** vlevo) podle směru, ve kterém se nachází další hraniční pixel. Po otestování všech pixelů tvořících hranice objektu a uložení směrovosti je Freemanův kód tvořen řetězcem čísel (**Obrázek 1** vpravo), které tvoří příznakový popis objektu. [2]



Obrázek 1: Osmi směrová reprezentace Freemanova kódu a Freemanův kód čtverce o straně délky 3 a začátkem v levém horním rohu

Rozpoznávání objektů (klasifikace) spočívá v zařazování objektů do předem daných tříd charakterizovaných jako množina prvků, které mají z hlediska klasifikace společné vlastnosti. Metod klasifikace je velké množství, od složitých neuronových sítí, které potřebují trénovací množinu dat, až po nejjednodušší porovnání vektorů příznaků s využitím minimální vzdálenosti, například Hammingova či Euklidovská vzdálenost. [3]

Klasifikace v této práci vychází z prostého porovnání řetězových kódů vstupního a idealizovaných objektů s velikostí a natočením odpovídající vstupnímu objektu. Vstupní objekt se porovná se všemi čtyřmi třídami (čtverec, obdélník, trojúhelník a kružnice) a klasifikuje se do té třídy, se kterou je nejvíce podobný. K porovnání lze využít například výše zmíněnou Hammingovu vzdálenost. Ukázka vstupního a výstupního objektu a jeho klasifikace (**Obrázek 2**).



Obrázek 2: Vstupní data (vlevo-čísla značí pořadí vzniku vektorů) a výstup (vpravo) - červeně značený idealizovaný objekt

3. HODNOCENÍ

Sada obrazců k testování vychází z psychology používaných kresebných testů, konkrétně z Bentonovy sady. Bentonova sada obsahuje obrazce složené z čtverců, obdélníků, trojúhelníků a kružnic. Pro správné určení třídy objektu jsou objekty klasifikovány jednotlivě dle čtyř elementárních tříd geometrických objektů, jak je zmíněno výše.

Hodnotí se celková odchylka kresby od předlohy a odchylky dílčích čar. Celková odchylka je míra nepřesnosti celého nakresleného objektu od předlohy, kdežto dílčí odchylka se týká každé čáry objektu zvlášť. Toto hodnocení je třeba kvůli hodnocení osobnostního rysu dané osoby. Může se totiž stát, že pokusný subjekt nakreslí jen část objektu s velkou odchylkou od předlohy. Může to být například proto, že subjektu dělá problém nakreslit čáry určitého směru nebo zakřivení.

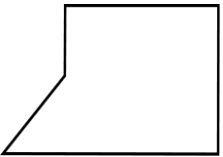
Dále se hodnotí míra odchylky pozice kresby vzhledem k podkladu. Protože administrace testu bude využívat jako kresebný podklad zařízení s dotykovým displejem, budou předlohy umístěny v různých pozicích kresebného okna. Je tedy třeba také zhodnotit, jestli subjekt svou kresbu umístil do stejné pozice, jako byla předloha.

ZÁVĚR

Freemanův kód byl aplikován na všechny čtyři základní primitiva a na testovaný obrazec. Testovaný tvar odpovídá čtverci s protaženým levým spodním rohem, což symbolizuje chybu při kresbě, kdy subjekt spodní hranu protáhne a pak se snaží vrátit se šikmým směrem vzhůru k počátku. Níže uvedený testovaný obrazec slouží pouze pro demonstraci metody a jako ukázka distorze, která může vzniknout při kreslení dle předlohy. V praxi je testováno mnohem více variant chyb každé třídy geometrických primitiv.

Na základě porovnání Hammingovy vzdálenosti Freemanových kódů byl testovaný obrazec (viz **Tabulka 1** **Chyba! Nenalezen zdroj odkazů.**) klasifikován jako čtverec, což dle tvaru testovaného objektu odpovídá. Druhý nejbližší je obdélník, naopak nejvzdálenější je trojúhelník s kružnicí. Navrženou metodou lze správně klasifikovat primitiva do tříd a rozšířit možnosti využití kresebných testů.

Tabulka 1: Hammingovy vzdálenosti testovacího tvaru a základních geometrických primitiv

Testovaný tvar	Freemanův kód	Objekt	Hammingova vzdálenost				
			Č	O	K	T	TT
	00000666664444422222	Čtverec (Č)	-	4	12	17	4
	0000000666444444222	Obdélník (O)	4	-	15	17	5
	076764646444222110	Kružnice (K)	12	15	-	12	12
	777777744444111111	Trojúhelník (T)	17	17	12	-	14
	0000066664444441122	Test. tvar (TT)	4	5	12	14	-

REFERENCE

- [1] HORÁK, Karel. *Popis objektů*. [cit. 2014-11-20]. Dostupné z: http://midas.uamt.feec.vutbr.cz/ZVS/lectures-pdf/10_Popis_objektu.pdf
- [2] ANNAPURNA, Pulipati, Sriraman KOTHURI a Srikanth LUKKA. *International Journal of Application or Innovation in Engineering* [online]. 2013 [cit. 2014-11-22]. ISBN 2319-4847. Dostupné z: <http://www.ijaiem.org/volume2issue8/IJAIEM-2013-08-31-081.pdf>
- [3] ŠŤASTNÝ, Jiří. *Netradiční metody a algoritmy pro rozpoznávání objektů technologické scény*. Brno: VUT FSI, 2005 [cit. 2014-11-27]. Dostupné z: <http://www.vutium.vutbr.cz/tituly/pdf/ukazka/80-214-3117-2.pdf>