

Sledování rtů v reálném čase pomocí Aktivních kontur

Miroslav Hlaváč¹

1 Úvod

V úloze audiovizuálního rozpoznávání řeči je důležitou informací tvar a poloha rtů řečníka. Tuto informaci je možné získat na základě různých metod detekce, založených například na segmentaci nebo modelování. Jelikož se tvar sledovaných rtů mění velmi rychle, je vhodné pro sledování využít metod méně výpočetně náročných, které jsou schopny konvergovat k řešení v desítkách milisekund.

2 Sledování rtů v reálném čase

Sledování rtů probíhá v sekvenci po sobě následujících obrazů. Při použití běžné kamery je potřeba zpracovat 25 snímků za vteřinu. Experimentálně jsem zjistil, že i při této rychlosti dochází k velkým změnám ve tvaru rtů mezi dvěma po sobě následujícími snímky. Velká změna se projeví ve vyšší výpočetní náročnosti v přechodu od tvaru detekovaném v prvním obraze do tvaru hledaném v obraze následujícím. Při hledání metody, která je schopná pracovat v reálném čase jsem narazil na metodu Aktivních kontur, kterou se dále zabývá tento článek.

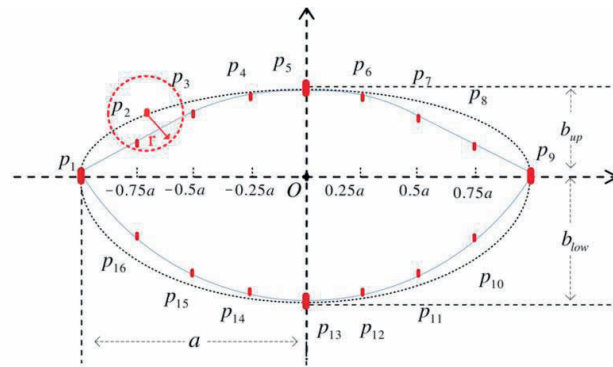
3 Aktivní kontury

Metoda Aktivních kontur (ACM) je spline minimalizující specifickou funkci globální energie. Počáteční kontura je deformována na základě omezující síly tak, aby detekovala hranice objektů. Základní algoritmus ACM je náchylný na intenzitu osvětlení, inicializaci parametrů a přítomnost zubů v obraze [1]. V článku [2] je prezentováno vylepšení metody lokalizace na základě rozdělení okolí na vnitřní a vnější, díky čemuž pak může být vypočtena potřebná lokální energie pro vývoj a extrakci. V článku [3] je dále zavedeno vylepšení metody o automatickou selekci parametrů modelu a nalezení minimální elipsy obalující počáteční objekt.

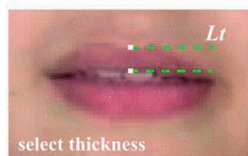
Po detekci tvaru rtů v prvním obraze je tento tvar následně použit jako výchozí informace pro hledání tvaru v obraze následujícím. Z detekovaného tvaru můžeme vypočítat šířku rtu (Obrázek 1.b) a použít ji k nastavení prohledávaného okolí, tak aby byl omezen vliv výskytu zubů a jazyka na přesnost detekce.

Tato metoda dosahuje až 95% úspěšnosti detekce hranice rtů. Případy selhání metody jsou většinou způsobeny přítomností vousů nebo nedostatečným kontrastem mezi barvou rtů a barvou okolní pokožky.

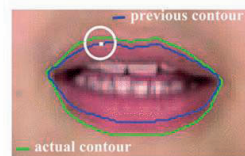
¹ student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, e-mail: mhlavac@kky.zcu.cz



(a) Initial evolving curve with the 16-point lip model



(b) Upper lip thickness



(c) Opening mouth

Obrázek 1: Správné parametry algoritmu [3]

4 Závěr

V článku byla představena metoda Aktivních kontur, kterou je možno použít v úloze sledování rtů v reálném čase. Informace o tvaru rtů lze dále použít k vylepšení přesnosti algoritmů rozpoznávání řeči v prostředí, kde díky šumu nebo postižení řečnicka selhávají běžné metody, založené pouze na rozpoznávání z audio signálu.

Poděkování

Příspěvek byl podpořen grantovým projektem SGS-2013-032 (Inteligentní metody strojového vnímání a porozumění 2)

Literatura

- [1] G. Chiou and J. N. Hwang, "Lipreading from color video," IEEE Transactions on Image Processing, vol. 6, no. 8, pp.1192–1195, 1997.
- [2] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," IEEE Transactions on Image Processing, vol. 17, no. 11, pp. 2029–2039, 2008.
- [3] Xin Liu; Yiu-ming Cheung; Meng Li; Hailin Liu, "A Lip Contour Extraction Method Using Localized Active Contour Model with Automatic Parameter Selection," Pattern Recognition (ICPR), 2010 20th International Conference on, pp. 4332,4335, 2010