

## Mixtura Gaussiánů

Mixtura Gaussiánů (MoG) je jednou z nejpoužívanějších rekurzivních technik, se kterou se lze v literatuře setkat. My zde popíšeme originální variantu algoritmu publikovanou Staufferem a Grimsonem, přičemž si tento text neklade za cíl precizní odvození veškeré související teorie.

Jak již název samotné metody MoG napovídá, k modelování jasů pixelů pozadí se používají distribuční funkce normálního rozdělení, které bývá také označováno jako Gaussovo rozdělení. Jednotlivé funkce hustoty pravděpodobnosti reprezentují možné typy povrchů vyskytujících se v místě daného pixelu. Takovýmto způsobem tedy pomocí funkce  $f$  modelujeme pravděpodobnost výskytu  $k$ -tého povrchu na určité pozici pixelu  $X$  v  $n$ -rozměrném (kanálovém) obraze.

$$f_{\mathbf{X}|k}(X|k, \Theta_k) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(X-\mu_k)^T \Sigma_k^{-1}(X-\mu_k)}, \quad (1.3)$$

kde  $\Theta_k = \{\mu_k, \sigma_k\}$  a diagonální kovarianční matice  $\Sigma_k = \sigma_k^2 I$ . Index  $k$  označuje povrch, který se v daném čase objevil na místě zpracovávaného pixelu a utváří jeho výslednou barvu. Ze zápisu kovarianční matice je zřejmé, že předpokládáme vzájemnou nezávislost hodnot v jednotlivých kanálech obrazu. To sice neodpovídá zcela skutečnosti, ale na výsledek funkce naší implementace to nebude mít zásadní vliv.

Pro usnadnění celé implementace je možné pracovat pouze s jednokanálovým obrazem ve stupních šedi. Funkce hustoty pravděpodobnosti poté nabývá jednodušší tvar

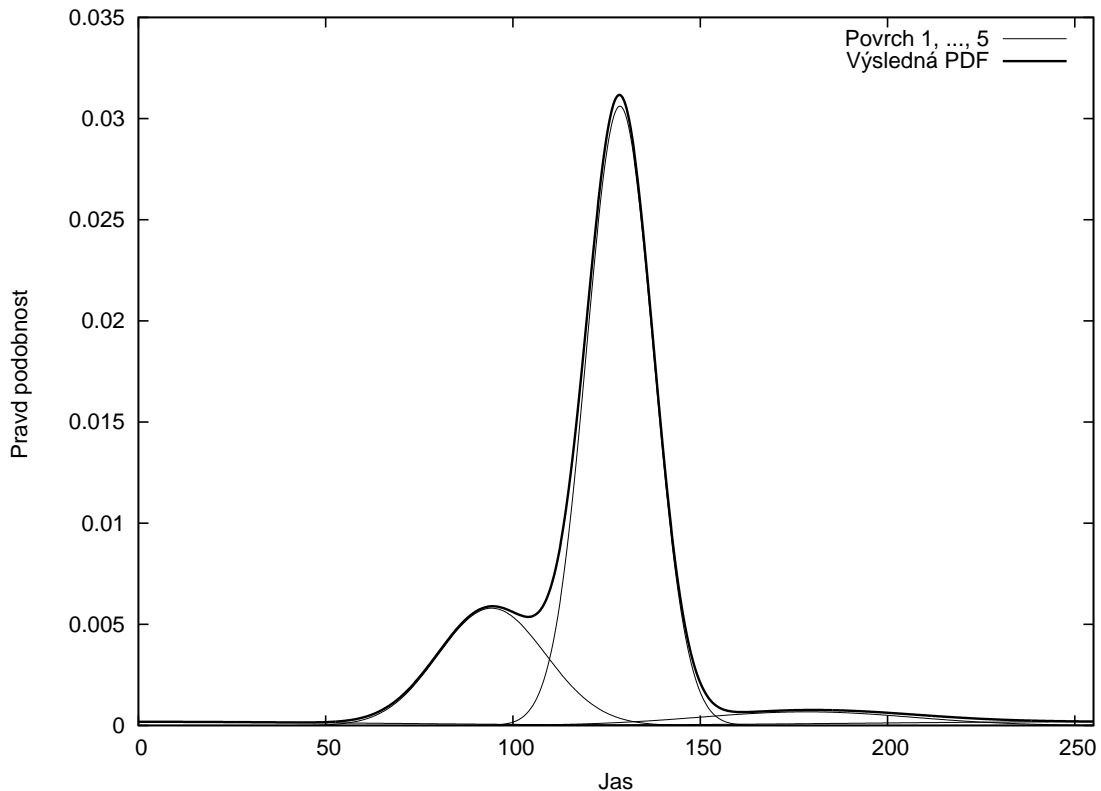
$$f_{\mathbf{X}|k}(X|k, \Theta_k) = \frac{1}{\sigma_k \sqrt{2\pi}} e^{\left(-\frac{(X-\mu_k)^2}{2\sigma_k^2}\right)}. \quad (1.4)$$

Jednotlivé funkce hustoty pravděpodobnosti jsou vyhodnocovány pro každý pixel obrazu zvlášť a jsou dynamicky aktualizovány pomocí změny korespondujících parametrů  $\mu_k$  a  $\sigma_k$ .

V našem případě použijeme 1-rozměrný šedotónový obraz (tj.  $X = (x)$  a  $X = (x_r, x_g, x_b)$  pro barevný) pořízený stacionární kamerou. Vzhledem k povaze popsaného modelu pozadí je důležité, aby bylo upevnění kamery dostatečně tuhé. V opačném případě by mohlo dojít třeba i k nepatrnému rozkmitání kamery a to by do značné míry zhoršovalo kvalitu detekce. V případě, že nemůžeme upevnění kamery ovlivnit, je možné této analýze předřadit vhodný stabilizační algoritmus.

V tomto okamžiku, kdy ke každému povrchu máme jednu hustotu pravděpodobnosti, potřebuje tyto funkce sloučit do jedné hustoty charakterizující konkrétní pixel. Jelikož je výskyt jednotlivých povrchů navzájem nezávislý, lze toto provést pomocí vážené sumy vyjádřené následující rovnicí

$$f_{\mathbf{X}}(X|\Phi) = \sum_{k=1}^K P(k) f_{\mathbf{X}|k}(X|k, \Theta_k), \quad (1.5)$$



Obrázek 1.2: Funkce hustoty pravděpodobnosti (ang. probability density function, zkráceně PDF) pro vybraný pixel šedotónového obrázku

kde  $P(k)$  představuje *a priori* pravděpodobnost výskytu  $k$ -tého povrchu na pozici analyzovaného pixelu a  $\Phi$  je množina parametrů  $\Theta_k$  pro všech  $k$  povrchů. Doplňme, že integrál přes celý rozsah hodnot  $X$  je roven 1 a totéž by mělo být dodrženo i pro diskrétní hodnoty.

Možný průběh funkce 1.5 lze vidět na obrázku 1.2. Čárkované čáry představují průběhy jednotlivých hustot pravděpodobností z rovnice 1.4, které jsou násobeny koeficientem pravděpodobnosti  $P(k)$ . V uvedeném příkladě je modelováno celkem 5 povrchů.

Z grafu 1.2 je dobře patrné, jak hodnoty kolem jasu 130 se v daném místě obrazu vyskytují nejčastěji a proto lze předpokládat, že právě Gaussián, který je v daném intervalu dominantní modeluje jas pozadí. Druhý výrazný Gaussián nalevo již může představovat popředí. Zbylé tři Gaussiány mají přiřazenu příliš malou pravděpodobnost a nejsou příliš patrné. Evidentně tedy modelují náhodné změny jasu pixelu, jež mohou být způsobeny např. projíždějícími automobily. Tato klasifikace, která má za úkol rozhodnout, zdali je Gaussián považován za model pozadí nebo popředí, je do značné míry volná. Můžeme se rozhodovat buď na základě pořadí Gaussiánů uspořádaných podle svého maxima nebo podle jejich tvaru, tj. parametru  $\sigma$ .

V dalším kroku potřebujeme získat výraz, který nám umožní vypočítat pravděpodobnost výskytu  $k$ -tého povrchu za podmínky, že známe aktuální hodnotu určitého pixelu ob-

razu. Požadovaný výraz obdržíme jako *a posteriorní* pravděpodobnost neboli podmíněnou pravděpodobnost s pomocí Bayesova teorému

$$P(k|X, \Phi) = \frac{P(k)f_{\mathbf{X}|k}(X|k, \Theta_k)}{f_{\mathbf{X}}(X|\Phi)}. \quad (1.6)$$

Nyní hledáme odhad, který z povrchů  $\hat{k}$  je s největší pravděpodobností na místě pixelu s aktuální hodnotou  $X$ . To lze vyjádřit pomocí následující rovnice

$$\hat{k} = \arg \max_k P(k)f_{\mathbf{X}|k}(X|k, \Theta_k). \quad (1.7)$$

Všimněme si, že v rovnici 1.7 je člen  $f_{\mathbf{X}}(X|\Phi)$  ze jmenovatele v rovnici 1.6 zcela vynechán. Důvod je zřejmý, jmenovatel je totiž nezávislý na  $k$  a nijak neovlivní výběr parametru  $\hat{k}$  maximalizující výraz 1.7.

Nyní se dostáváme k aktualizaci parametrů  $P(k)$ ,  $\mu_k$  a  $\sigma_k$  charakterizujících náš model. K tomu použijeme obvyklý způsob, který bývá označován jako přímé průměrování (ang. on-line averages). Výpočet provedeme podle následujícího schématu

$$\hat{P}(k) = (1 - \alpha)P(k) + \alpha P(k|X, \Phi) \quad (1.8)$$

$$\hat{\mu}_k = (1 - \rho_k)\mu_k + \rho_k X \quad (1.9)$$

a

$$\hat{\sigma}_k^2 = (1 - \rho_k)\sigma_k^2 + \rho_k ((X - \hat{\mu}_k) \circ (X - \hat{\mu}_k)), \quad (1.10)$$

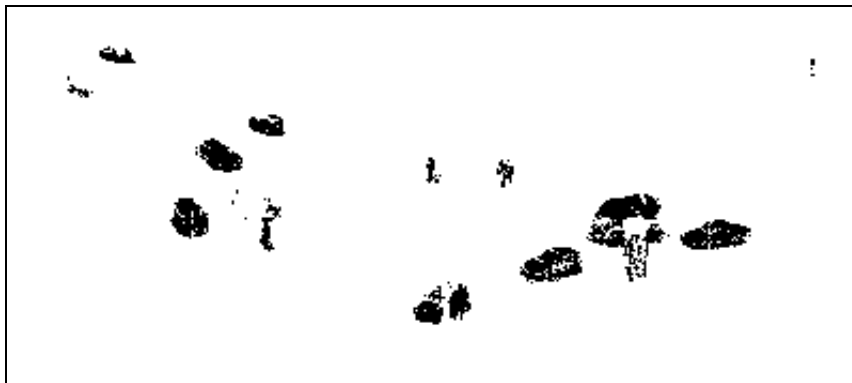
kde  $\rho_k = \alpha P(k|X, \Phi)/\hat{P}_k$  a nově vypočtené hodnoty parametrů jsou označeny symbolem stříšky.

Závěrečným krokem je rozlišit jednotlivé povrchy na pozadí a popředí. Je možno experimentovat s několika přístupy, které již byly naznačeny během popisu grafu 1.2. Nejjednodušší přístup je zvolit určitý práh, po jehož překročení bude každý Gaussián a jemu přiřazený  $k$ -tý povrch odpovídat pozadí. Důležité je zachovat možnost zachytit multimodální pozadí, tzn., aby mohlo být pozadí reprezentováno více jak jedním povrchem. Výsledek činnosti popsaného algoritmu si je možno prohlédnout na obrázku 1.3.

## Detekce popředí

Cílem detekce popředí je identifikovat pixely, jejichž hodnota jasů neodpovídá aktuálnímu stavu modelu. Model je u většiny zmíněných metod prakticky tvořen pouze jedním obrazem  $b_t$  (mimo neparаметrický model a mixturu Gaussiánů). Nejjednodušší metodu detekce popředí lze zapsat ve tvaru výrazu 1.1 takto

$$g(x, y) = |f_t(x, y) - b_t(x, y)|. \quad (1.11)$$



Obrázek 1.3: Výstup metody MoG: (shora dolů) původní snímek z video sekvence, odpovídající maska zachycující výslednou segmentaci

Uvedený vztah 1.11 je možno normalizovat pomocí statistických charakteristik modelu  $\mu_d$  a  $\sigma_d$ , které získáme z výrazu  $f_t(x, y) - b_t(x, y)$  vypočteném ve všech bodech obrazu. Pozměněná funkce  $g$  pak nabude tvar

$$g(x, y) = \frac{|f_t(x, y) - b_t(x, y) - \mu_d|}{\sigma_d}. \quad (1.12)$$

Výsledkem takovéto úpravy je vyšší stabilita algoritmu. Zbývá již jen experimentálně určit parametr  $\tau$ . Na závěr lze ještě zmínit možnost použít místo absolutní difference ve vztahu 1.11 diferencí relativní

$$g(x, y) = \frac{|f_t(x, y) - b_t(x, y)|}{b_t(x, y)}, \quad (1.13)$$

případně lokalizovat platnost prahu  $\tau$  do nějaké podoblasti celkového obrazu. Pak by každé takovéto oblasti  $\omega_i$  náležel jeden parametr  $\tau_i$ , což by vedlo k navýšení kontrastu ve tmavých částech obrazu.

Sofistikovanější detekce je založena na metodě MoG a detekce popředí vychází z postupu popsaného v závěru rozboru této metody. Stačí tedy v masce  $m$  nastavit pixely na hodnotu 1 tam, kde aktuální hodnota pixelu nemůže být modelem klasifikována jako pozadí.

### 1.1.5 Validace dat

Z praktických pokusů s výše uvedenými metodami jednoznačně vyplývá, že výsledná maska popředí není v řadě případů zcela spolehlivá. Mezi příčiny tohoto stavu patří ignorace přirozené korelace mezi sousedními pixely, rychlost adaptace modelu nemusí být v souladu s rychlostí pohybujících se objektů a např. mohou být oblasti zastíněné pohybujícím se objektem chybně označeny za popředí.

První uvedený problém projevující se ve formě nesouvislosti masky v aktuálním místě pohybujícího se předmětu lze poměrně úspěšně řešit pomocí morfologických operátorů a algoritmu pro shlukování společných komponent. Výsledné komponenty s příliš malou plochou lze z masky odstranit.

Častým problémem je také zanechávání stop za objektem nebo vytváření děr v objektu. Tento jev se nejčastěji vyskytuje v případě, že se obrazem pohybuje podlouhlý objekt. Budeme-li za sledovaný objekt uvažovat homogenní vlakovou soupravu, pak lokomotiva bude označena korektně. S postupnou adaptací modelu na novou texturu začne docházet od určitého místa naší soupravy k mizení vagonů. Ty budou totiž chybně zařazeny do modelu pozadí. Po přejetí posledního vagonu budou pixely skutečného pozadí chybně vyhodnoceny jako pixely popředí a vznikne tzv. duch. Po uplynutí určité doby se model opět přenastaví a bude za pozadí považovat opět správně jasy skutečného pozadí a stopa vymizí. Jednoduchým řešením takovýchto problémů je využití vícero modelů pozadí aktualizovaných s rozdílnou frekvencí a výslednou segmentaci založit na porovnávání výsledků získaných ze separátních modelů.