هفتمین کنفرانس ملی مهندسی برق و الکترونیک ایران
دانشگاه آزاد اسلامی گناباد
۲۸ و ۲۹ مرداد ماه ۹۴

بخش بندی تصاویر براساس مدل میدان تصادفی مخفی مارکوف

سید مهدی صفوی[۱] ، محمد میکاییلی[۲]، محمد پویان[۳]

۱-دانشجو گروه مهندسی پزشکی دانشکده فنی مهندسی دانشگاه شاهد

۲و۳- عضوهیات علمی دانشیار گروه مهندسی پزشکی دانشکده فنی مهندسی دانشگاه شاهد

safavi.mehdi@gmail.com
pooyan@shahed.ac.ir
m_mikaili@yahoo.com

***چکیده* — ما با مطالعه مدل مخفی میدان تصادفی مارکوف از الگوریتم EM جهت تخمین پارامترهای مدل استفاده کردیم تا بتوانیم بر اساس آن بخش بندی تصویر را انجام دهیم. در این مقاله برای پارامترها از مدل گوسی استفاده شده که طی الگوریتم EM تخمین زده شده و سپس با روش MAP بخش بندی تصویر بر اساس تئوری Hammersly-Cliford بروزرسانی میشود. پیاده سازی مدل در بخشهای مجزا ارایه شده است.**

واژه‌های کلیدی* — مدل مخفی میدان تصادفی مارکوف (HMRF)Hidden Markov Random Field Model ، EM (Expectation Maximization)، MAP (Maximum Aposterior Probability)

۱.مقدمه:

میدان های تصادفی مارکوف (MRF) در مسایل بینایی ماشین مانند بخش بندی تصویر و بازسازی آن کاربرد دارد. از ویژگیهای بارز آن میتوان به الگوریتمهای کارا، مانند مدهای شرطی تکراری با لحاظ وفاداری به دیتا و همواری و نرمی مدل اشاره نمود.

قالب HMRF-EM در ابتدا برای بخشبندی تصاویر MRI ارایه شد[۷]

در یک تصویر متشکل از N پیکسل، شدت روشنایی تصویر را با بردار$Y = (y_0, y_1, y_2, \dots, y_N)$ نشان میدهیم که y_i شدت روشنایی پیکسل i ام میباشد. در تصویر رنگی شدت روشنایی در سه بعد R، G و B بصورت مجزا YR، YG و YB بیان میشود.

حاصل بخش بندی تصویر، یک تصویر معادل برچسب گذاری شده است که با بردار$X = (x_0, x_1, x_2, \dots, x_N)$ که در آن $x_i \in L$ بیان میشود. برای مثال در بخش بندی باینری مجموعه ی L دارای دو عضو $L = \{0,1\}$ بوده و x_i مقدار 0 یا 1 را خواهد داشت.

برای برچسب گذاری تصویر، طبق معیار MAP در جستجوی X^* ای هستیم که حداکثر احتمال X را با دانستن Y ، طبق رابطه (۱) برآورده کند.

$$X^* = argmax\,\{P(X|Y)\} = \\ = argmax\,\{P(Y|X,\Theta)P(X)\} \qquad (1)$$

که در آن$P(X)$ احتمال پیشین با توزیع Gibbs و $P(Y|X,\Theta)$ احتمال توام درست نمایی از رابطه (۲) بدست می

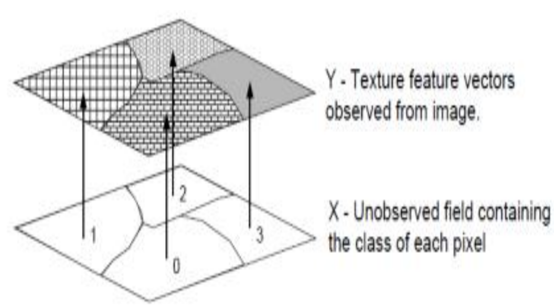
آید که در آن $P(y_i|x_i,\theta_{x_i})$ یک توزیع گوسی با پارامترهای $\theta_{x_i}=(\mu_{x_i},\sigma_{x_i})$ است.

$$P(Y|X,\Theta)=\prod_i P(y_i|X,\Theta)$$
$$=\prod_i P(y_i\mid x_i,\theta_{x_i}) \qquad (۲)$$

معمولا در مسائل MRF ، آموزش مجموعه پارامتر θ از طریق دیتای آموزشی صورت میگیرد. برای مثال در مسائل بخش بندی دانش قبلی راجع به توزیع شدت پس زمینه و روی زمینه ممکن است در یک مجموعه اطلاعات (دیتا ست) ثابت باشد. بنابراین ما می توانیم پارامتراها را از طریق برخی تصاویر که بصورت دستی برچسب گذاری شده آموزش دهیم و آن پارامترها را برای اجرای MRF جهت بخشبندی سایر تصاویر بکاربندیم.

تفاوت اصلی بین MRF و HMRF اینست که در HMRFمجموعه پارامترهای θ بدون راهنما آموزش داده میشوند (unsupervised).

در مدل HMRF برای مساله بخشبندی تصویر، نیازی به مرحله آموزش نیست و فرض میشود دانشی راجع به توزیع شدت روشنایی پیش زمینه و پس زمینه ارایه نشده باشد. بنابراین الگوریتم EM جهت بدست آوردن مجموعه پارامتر $\Theta=\{\theta_l\mid l\in L\}$ پیشنهاد میشود که در سیر تکراری اطلاعات آن تکمیل میشود.



شکل(۱)

۲.مواد و روشها:

تئوری Hammersly-Cliford

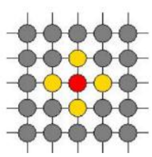
در تصاویر حقیقی پیکسلهای همسایگی معمولا ویژگی مشابهی دارند (شدت روشنایی، رنگ و بافت و ...) که براساس این تئوری اگر $P(X)$ یک حوزه برچسب زنی MRF باشد آنگاه بایستی از توزیع Gibbs پیروی کند. پس احتمال تخصیص برچسب به هر پیکسل به شرط مشاهده همسایه ها ست.

$$P(X)=\frac{1}{Z}exp\big(-U(X)\big)$$
$$U(X)=\sum_{\substack{j\in N_i\\ c\in C}} V_c(l,x_i)$$

تعریف کلیک یا دسته Clique

زیر مجموعه C از همسایگی هر پیکسل را در حوزه برچسب زنی یک دسته میگویند اگر هر جفت پیکسل در این زیر مجموعه همسایه باشند. این همسایگی میتواند هشت تایی و چهارتایی تعریف شود. در این مقاله برای هر پیکسل دسته چهار همسایگی، مطابق شکل در نظر گرفته شده است.

1	1	1	1	3	3	3
1	1	1	1	1	1	1
1	1	2	2	2	2	1
1	1	2	2	2	2	1
1	1	1	1	1	1	1



شکل(۲)

برای محاسبه انرژی در حوزه لیبل تصویرELF ، به لیبلهای کلیک یا دسته ی هر پیکسل توجه میکنیم .

$$ELF=U(X)$$

برای تمام پیکسهای تصویر میزان شباهت لیبل دسته ها با لیبل l ام تصویر با $I_{x_i,l}$ تعیین میشود.

$$I_{x_i,l}=\begin{cases}0 & if \quad x_i\neq l\\ 1 & if \quad x_i=l\end{cases}$$

پس از شباهت سنجی، میتوان پتانسیل هر دسته را طبق رابطه () بدست آورد. و یا با رابطه () مستقیما پتانسیل هر دسته را در حوزه لیبل l ام تصویر بدست آورد.

$$V_c(l,x_i)=\frac{1}{2}(1-I_{x_i,l})$$
$$V_c(l,x_i)=\begin{cases}\frac{1}{2} & if \quad x_i\neq l\\ 0 & if \quad x_i=l\end{cases}$$

بنابراین انرژی هر دسته در حوزه لیبلl ام ELF قابل محاسبه خواهد بود.

الگوریتم EM

ما از الگوریتم EM برای تخمین مجموعه پارامتر Θ طبق مراحل زیر استفاده کرده ایم.

شروع: فرض کنید مجموعه پارامتر اولیه ما $\Theta^{(0)}$ باشد.

مرحلهE : $\Theta^{(t)}$ مجموعه پارامتر در t امین تکرار بوده و امید شرطی را طبق رابطه زیر بدست می آوریم

$$Q(\Theta|\Theta^{(t)}) = E[ln\, P(X,Y|\Theta)|Y,\Theta^{(t)}] = \sum_{x\in\chi}\left[P(X|Y,\Theta)|\ln P(X,Y|\Theta^{(t)})\right] \qquad (۳)$$

که χ مجموعه ای از مقادیر ممکن برچسب هاست.

مرحلهM : حالا بایستی $Q(\Theta|\Theta^{(t)})$ را بیشینه کنیم تا $\Theta^{(t+1)}$ تخمین بعدی مجموعه پارامتر را بدست آوریم

$$\Theta^{(t+1)} = \arg\max_{\Theta}\{Q(\Theta|\Theta^{(t)})\} \qquad (٤)$$

پس با جاگذاری $\Theta^{(t+1)} \longrightarrow \Theta^{(t)}$ مرحله E را تکرار می کنیم.

فرضیات مساله :

با یک سری از فرضیات به تشریح الگوریتم خواهیم پرداخت:

فرض۱: اگر $G(z;\theta_l)$ تابع توزیع گوسی با پارامترهای $\theta_l=(\mu_l,\sigma_l)$ باشد:

$$G(z;\theta_l)=\frac{1}{\sqrt{2\pi}\sigma_l}\exp\left(-\frac{(z-\mu_l)^2}{2\sigma_l^2}\right) \qquad (٥)$$

فرض۲: احتمال پیشین $P(X)$ را به صورت یک توزیع Gibbs بیان می‌کنیم که در آن $U(X)$ را تابع انرژی حوزه لیبل می نامیم.

$$P(X)=\frac{1}{Z}\exp(-U(X)) \qquad (٦)$$

فرض۳ : همچنین اگر $P(Y|X,\Theta)$ نیز از توزیع گوسی حاصل شده باشد آنگاه میتوان آن را در فرم توزیع Gibbs طبق رابطه زیر بیان نمود

$$P(Y|X,\Theta)=\prod_i P(y_i|x_i,\theta_{x_i}) = \prod_i G(y_i;\theta_{x_i}) = \frac{1}{Z'}\exp(-U(Y|X)) = \prod_i \frac{1}{\sqrt{2\pi}\sigma_l}\exp\left(-\frac{(y_i-\mu_l)^2}{2\sigma_l^2}\right) = \frac{1}{Z'}\exp(-U(Y|X)) \qquad (٧)$$

محاسبه مجموعه پارامتر اولیه با k-means

برای مدلسازی بر پایه توزیع گوسی بایستی یک بخش بندی اولیه داشته باشیم . این بخش بندی اولیه یا به عنوان اطلاعات اولیه داده میشود یا ازطریق الگوریتمی آن را بدست آورد.

در این بخش با استفاده از الگوریتم خوشه بندی k-means یک بخش بندی اولیه روی شدت سطح خاکستری پیکسل های تصویر اعمال میکنیم تا برچسب اولیه $X^{(0)}$ را در k کلاس فراهم شود. حال می توان پارامترهای اولیه مدل $\Theta^{(0)}$ را برای هر کلاس بدست آورد. $\boldsymbol{\theta_l = (\mu_l, \sigma_l)}$

$$\Theta=\{\theta_l \mid l\in L\}$$

الگوریتم EM

طبق مراحل زیر تغییر خواهد کرد

شروع با مجموعه پارامتر اولیه $\Theta^{(0)}$

مدلسازی توزیع روی هر لیبل و محاسبه انرژی حوزه برچسب گذاری EFF و انرژی حوزه ویژگیELF(شدت روشنایی) آن

$$\boldsymbol{EFF_l = U(Y|X,\,\Theta_l) = \sum_i U(y_i|x_i,\Theta_l) = \sum_i \frac{(y_i-\mu_{x_i})^2}{2\sigma_{x_i}^{\,2}} + \ln\sigma_{x_i}}$$

$$\boldsymbol{ELF_l = U(X) = \sum_{j\in N_i} V_c(l, x_j^{(k)})}$$

با استفاده از مجموعه پارامترفعلی $\Theta^{(t)}$ ، $X^{(t)}$ را آپدیت کنید. که با روشMAP تخمین $X^{(t)}$ انجام میشود. در بخش۵ بحث خواهد شد.

$$X^{(t)}=\arg\max_{X\in\chi}\{P^{(t)}(Y|X,\Theta^{(t)}).P(X)\} = \arg\min_{X\in\chi}\{U(Y|X,\Theta^{(t)})+U(X)\} \qquad (٨)$$

محاسبه چگالی احتمال توزیع پسین برای تمام $l\in L$ و y_i که $x_{N_i}^{(t)}$ پیکره بندی همسایگی $x_i^{(t)}$

$$P^{(t)}(l|y_i)=\frac{G(y_i;\theta_l)P(l|x_{N_i}^{(t)})}{P^{(t)}(y_i)} \qquad (٩)$$

$$P^{(t)}(y_i)=\sum_{l\in L}G(y_i;\theta_l)P(l|x_{N_i}^{(t)}) \qquad (١٠)$$

$$P(l|x_{N_i}^{(t)})=\frac{1}{Z}\exp\left(-\sum_{j\in N_i}V_c(l,x_j^{(t)})\right) \qquad (١١)$$

۱. استفاده از $P^{(t)}(l|y_i)$ برای آپدیت پارامترهای مدل

$$\mu_l^{(t+1)} = \frac{\sum_i P^{(t)}(l \mid y_i).y_i}{\sum_i P^{(t)}(l \mid y_i)} \quad (۱۲)$$

$$(\sigma_l^{(t+1)})^2 = \frac{\sum_i P^{(t)}(l \mid y_i).(y_i - \mu_l^{(t+1)})^2}{\sum_i P^{(t)}(l \mid y_i)} \quad (۱۳)$$

۲. بررسی شروط همگرایی EFF+ELF یا حداکثر تکرار الگوریتم جهت اتمام و یا ادامه الگوریتم که با جاگذاری $\Theta^{(t+1)} \longrightarrow \Theta^{(t)}$ و پرش به مرحله ۲ میسر خواهد بود.

تخمین MAP

با توجه به اینکه درالگوریتم EM نیاز به یافتن X^* است که در آن کل انرژی پسین حداقل می‌شود

$$X^* = \arg\min_{X \in \chi} \{U(Y \mid X,\Theta) + U(X)\} \quad (۱٤)$$

$$U(Y \mid X,\Theta) = \sum_i U(y_i \mid x_i,\Theta)$$

$$U(y_i \mid x_i,\Theta) = \frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^{\ 2}} + \ln \sigma_{x_i} \quad (۱۵)$$

$$U(Y \mid X,\Theta) = \sum_i \left(\frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^{\ 2}} + \ln \sigma_{x_i} \right)$$

$$U(X) = \sum_{c \in C} V_c(X) \quad (۱٦)$$

که $V_c(X)$ پتانسیل دسته یا گروه همسایگی و C مجموعه تمام دسته های ممکن می‌باشد.

ما در حوزه تصویر فرض می‌کنیم که یک پیکسل دارای حداکثر چهار همسایگی است بنابراین پتانسیل گروه بر روی جفتهای همسایگی پیکسلها تعریف میشود

$$V_c(x_i, x_j) = \frac{1}{2}(1 - I_{x_i,x_j}) \quad (۱۷)$$

$$I_{x_i,x_j} = \begin{cases} 0 & if\ x_i \neq x_j \\ 1 & if\ x_i = x_j \end{cases} \quad (۱۸)$$

ما الگوریتم تکراری را برای حل (۱٤)توسعه دادیم با:

شروع راه حل با تخمین اولیه‌ی $X^{(0)}$ است که از حلقه قبلی الگوریتم EM گرفته شده است

و جستجو می‌کنیم $X^{(k)}$ آماده شده برای $1 < i < N$

$$X^{(k+1)} = \underset{l \in L}{\operatorname{argmin}} \{ U(y_i \mid l) + \sum_{j \in N_i} V_c(l, x_j^{(k)}) \}$$

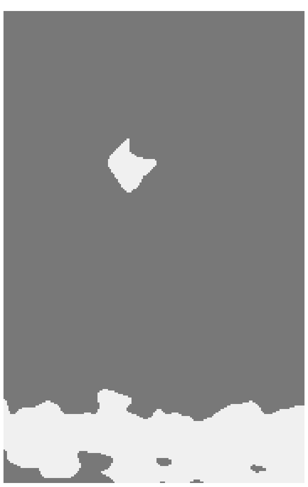
تکرار مرحله ۲ تا همگرایی $U(Y|X,\Theta) + U(X)$ یا رسیدن به حداکثر تکرارk

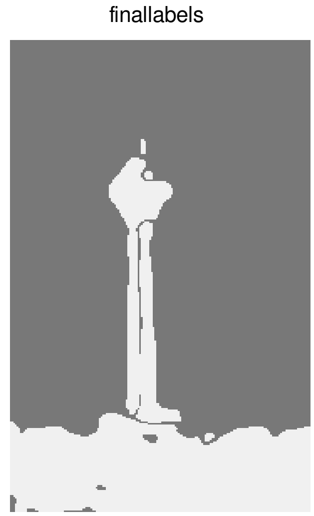
۳.جمع بندی :

با الگوریتم EM بخش بندی تصاویر در میدان مدل مخفی تصادفی مارکوف پیاده سازی انجام شد. این الگوریتم بر روی تصویر حاصل از روش kmeans پالایش انجام میدهد. در تصاویری که پیچیدگی دارند در تخمین توزیع شدت انرژی پیکسل های آن، نمیتوان نتایج پالایشی خوبی گرفت و باید از ویژگی دیگری برای آن استفاده کرد. نتیجه اعمال الگوریتم ارایه شده در شکل(۳) ارایه شده است.

Input Image

initial labels from kmeans





شکل(۳)