



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - ۵ و ۶ اسفند ماه ۱۳۸۸

تشخیص و ردیابی بی درنگ چهره در مقابل عکس و سایر اشیاء مبتنی بر PCA و رنگ پوست و ویژگی های SMQT با استفاده از فیلتر ذره ای

فاطمه حورعلی^(۱) - محمدحسین صدیقی^(۲)

(۱) گروه برق - دانشگاه صنعتی سهند تبریز

Fhourali@yahoo.com

(۲) گروه برق - دانشگاه صنعتی سهند تبریز

Sedaaghi@sut.ac.ir

خلاصه: تشخیص و ردیابی چهره کاربردهایی در حوزه هایی مانند ساختارهای ویدئویی و شاخص گذاری و نظارت بصری (موارد امنیتی) و ایجاد حوزه های تحقیقاتی، دارد. این مقاله الگوریتمی مقاوم را برای تشخیص و ردیابی بی درنگ چهره انسان در مقابل عکس و دیگر اشیاء با وجود شرایط نوری، حالات و زوایای چهره مختلف ارائه می دهد. این الگوریتم بر پایه رنگ پوست، ویژگی های SMQT، PCA و استفاده از اطلاعات حرکتی و فیلتر ذره ای برای ردیابی می باشد. در مرحله اول تشخیص چهره با استفاده از بخش بندی نواحی رنگ پوست و ویژگی های SMQT انجام می شود. ردیابی براساس یادگیری افزایشی یک نمایش زیرفضای چهره با بعد کم، تشخیص چهره و استفاده از فیلتر ذره ای می باشد. روش ردیابی و تشخیص چهره پیشنهاد شده عملکرد بالا و پایداری در مقابل تغییرات مقیاس، چرخش سر و همچنین حرکات سریع شخص را نتیجه می دهد، ضمن اینکه پیچیدگی محاسبات نیز نسبت به سایر روش ها به میزان زیادی کاهش می یابد. نتایج آزمایشات موثر بودن این الگوریتم و برتری آن نسبت به الگوریتم جابجایی میانگین، الگوریتم جابجایی میانگین بسط یافته براساس هیستوگرام و همچنین استفاده از فیلتر کالمن برای ردیابی چهره را نشان می دهد.

کلمات کلیدی: تشخیص چهره، ویژگی های SMQT، یادگیری زیرفضای چهره، فیلتر ذره ای، ردیابی چهره.

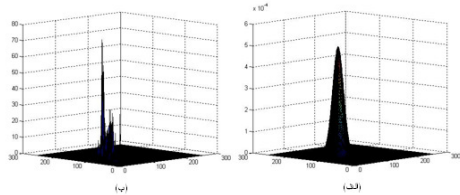
۱ - مقدمه

مسئله تشخیص و ردیابی چهره امری چالش انگیز می باشد زیرا چهره، هدفی غیرصلب می باشد که درجه بالایی از تغییرات نسبت به چرخش سر، روشنایی، حالات چهره و انسداد را شامل می شود. تشخیص و ردیابی اتوماتیک چهره امری مهم در بسیاری از کاربردهای تجاری مانند کنفرانس های ویدئویی، کنترل دستیابی امنیتی و شاخص گذاری ویدئوی مبتنی بر محتوا می باشد. بنابراین چندین تکنیک تشخیص و ردیابی چهره در چند دهه اخیر مورد مطالعه قرار گرفته است.

در این زمینه Suandi و سایرین تکنیکی را برای تخمین موقعیت چهره از توالی ویدئوی رنگی با استفاده از شبکه بیزین دینامیک (DBN) ارائه کردند. به دلیل اینکه ردیاب های چهره ویژگی های چهره مانند چشم، مردمک، گوشه های دهان و نواحی پوست را ردیابی می کنند، روش پیشنهادی آن ها صرفاً از ویژگی های مردمک ها و مرکز دهان و نواحی پوست برای محاسبه ملاک ها برای استنباط

DBN استفاده می کند [۱]. الگوریتمی برای ردیابی چهره و ویژگی های چهره براساس استخراج زاویه انحراف سر مبتنی بر ترکیبی از بخش بندی رنگ و برجسب زنی اجزای متصل به هم توسط Kabakli و سایرین ارائه شد [۲]. Bhandarkar و Zheng در [۳] الگوریتمی را که فیلتر ذره ای تطبیقی نامیده شده است برای تشخیص و ردیابی چهره استفاده کردند. الگوریتم پیشنهادی آن ها براساس ایجاد یک فیلتر ذره ای تطبیقی و الگوریتم تشخیص چهره Adaboost می باشد. Verma و سایرین [۴] یک روش احتمالاتی جدید برای تشخیص و ردیابی چهره در یک توالی ویدئویی را ارائه دادند. Wang و Liu یک روند جدید برای ردیابی و تشخیص چهره ترکیب شده در ویدئو را پیشنهاد کردند. الگوریتم تشخیص چهره آن ها یک روش تطبیقی کلیشه سریع با استفاده از برنامه سازی پویا می باشد [۵]. Vasant در [۶] یک الگوریتم تشخیص را پیشنهاد کرد که از یک روند آماری که توسط Viola و Jones ارائه شد [۷]، استفاده می کند. این الگوریتم از یک مجموعه ویژگی Haar-like و یک مجموعه پشت سرهم از

احتمال پوست نرمالیزه می شوند و به صورت یک تصویر سطح خاکستری نشان داده می شود. برای تبدیل این تصویر به یک تصویر باینری لازم است یک ترشلد تعیین شود. میزان ترشلد بهینه با یافتن نقطه ای که میزان تغییرات در نواحی بخش بندی شده مینیمم باشد، تعیین می شود. پس از تعیین مقدار ترشلد تمام مقادیر پیکسل هایی که احتمالی بالاتر از ترشلد دارند به یک و بقیه پیکسل ها به صفر نسبت داده می شوند. در نهایت عملگرهای مورفولوژیکی به منظور جدا کردن نواحی ای که با احتمال بیشتری مربوط به رنگ پوست می باشند اعمال می شوند [9]. شکل (۲) تمامی این مراحل را برای یک تصویر نمونه نشان می دهد.



شکل (۱): (الف) مدل گوسین دوبعدی (ب) هیستوگرام دوبعدی مقادیر از تصاویر نمونه پوست



شکل (۲): الف. تصویر رنگی اصلی ب. نواحی احتمالاتی پوست ج. نواحی بخش بندی شده پوست د. نواحی بخش بندی شده پوست پس اعمال عملگرهای مورفولوژیکی

سپس ویژگی های تبدیل کوانتیزه کردن میانگین متوالی برای استخراج ویژگی های غیرحساس به روشنایی بکار می روند. در واقع برای سرعت دادن به عملکرد تشخیص و همچنین کاهش انجام محاسبات، محاسبات مربوط به استخراج ویژگی های SMQT تنها در نواحی بخش بندی شده به رنگ پوست اعمال می شود. هم چنین یک شبکه جداساز غیرپیکسلی از تقسیمات برای سرعت دادن به کلاسیفایر اصلی استفاده می شود. در نهایت ویژگی ها و طبقه بند برای تشخیص چهره تمام رخ مورد استفاده قرار می گیرند. در واقع بسطی از طبقه بند SNOW ، SNOW جداساز ، برای طبقه بندی مورد استفاده قرار می گیرند. SNOW جداساز، نتایج کلاسیفایر SNOW اصلی را مورد استفاده قرار می دهد و یک توالی از طبقه بند ها را برای انجام تشخیص سریع تر ایجاد می کند.

SQMT روندی را استفاده می کند که یک تفکیک ساختاری اتوماتیک از اطلاعات انجام می دهد. این ویژگی ها می توانند روی نواحی محلی در یک تصویر برای استخراج ویژگی های غیرحساس به روشنایی بکار می روند. برای بررسی محلی، یکی از روش ها تقسیم بندی تصویر به بلوک هایی با یک سایز از قبل تعریف شده می باشد. فرض کنیم X یک پیکسل و $D(x)$ مجموعه ای از $|D(x)| = D$ پیکسل از یک ناحیه محلی در یک تصویر باشد. تبدیل SQMT از ناحیه محلی به صورت معادله (۱) می باشد.

$$SQMT_L : D(x) \rightarrow M(x) \quad (1)$$

طبقه بندهای درخت تصمیم گیری به عنوان یک مدل آماری استفاده می کند. هم چنین ردیابی به عنوان تشخیص پیوسته در دو فریم نگاشت داده شده با استفاده از یک الگوریتم بر پایه فاصله بین مراکز مستطیل های دربرگیرنده چهره انجام می شود. Liu و Yang یک روش تشخیص چهره سلسله مراتبی را با استفاده از تطبیق کلیشه و الگوریتم PCA دوبعدی ارائه دادند. این روش شامل ۲ کلاسیفایر مختلف می باشد. اولین طبقه بند بیشتر بخش های غیرچهره را فیلتر می کند و دومین طبقه بند از الگوریتم PCA دو بعدی برای تشخیص چهره براساس نتایج اولین طبقه بند استفاده می کند [8].

الگوریتم پیشنهاد شده در این مقاله شامل دو بخش می باشد: بخش ۱ تشخیص و بخش ۲ پیش بینی و به روز کردن مدل جهت ردیابی. در واقع ابتدا چهره های موجود در تصویر یا اولین فریم در یک توالی ویدئویی به کمک آشکارساز چهره بر مبنای ویژگی های SMQT و بخش بندی نواحی رنگ پوست تشخیص داده شده و با استفاده از اطلاعات حرکتی چهره انسان از چهره موجود در عکس تشخیص داده می شود. سپس در فریم های بعد با استفاده از یادگیری زیرفضای با بعد کاهش یافته چهره و تشخیص چهره و فیلتر ذره ای، ردیابی چهره تشخیص داده شده انجام می شود. مزیت این روش غیرحساس بودن به انسداد چهره، چرخش، تغییر حالت و تغییر مقیاس چهره و تغییر شرایط نوری می باشد. مزیت های دیگر الگوریتم پیشنهادی عبارتند از: (۱) توانایی تشخیص موقعیت اولیه چهره انسان در مقابل چهره موجود در عکس و سایر اشیاء (۲) غیرحساس بودن نسبت به تغییر در مقیاس و جهت چهره (۳) عدم تأثیر تغییرات شرایط نوری بر امر تشخیص و ردیابی. علاوه بر آن، این الگوریتم از لحاظ محاسباتی ساده و به صورت بدون درنگ کار می کند.

۲- الگوریتم پیشنهادی

الف- تشخیص چهره

الگوریتم پیشنهادی تشخیص چهره بر اساس تشخیص نواحی پوست و استفاده از ویژگی های SMQT می باشد.

مرحله اول بخش بندی پوست، ایجاد یک مدل پوست می باشد. تصویر ورودی به فرمت RGB و حساس به شرایط نوری می باشد. در اینجا سیستم رنگی YCbCr که در آن Y نشان دهنده مولفه روشنایی است درحالی که مولفه های Cr و Cb نشان دهنده مولفه های رنگ تابی یک تصویر رنگی می باشند، مورد استفاده قرار گرفته است. توزیع رنگ پوست افراد مختلف می تواند در یک ناحیه کوچک از فضای رنگی خوشه بندی شود. مولفه Y مورد بررسی قرار نمی گیرد زیرا به روشنایی مربوط است اما مولفه های Cr و Cb استفاده می شوند زیرا حاوی اطلاعات رنگی می باشند. برای تعیین توزیع رنگ پوست در فضای رنگی از هزاران پیکسل از نمونه های انتخاب شده از پوست افراد مختلف استفاده می شود. مدل گوسی دو بعدی و هیستوگرام دو بعدی مقادیر $CrCb$ مربوط به تصاویر پوست در شکل (۱) آمده است. پس از این که مدل پوست ایجاد شد، مقادیر احتمال برای هر پیکسل با محاسبه فاصله اقلیدسی از میانگین محاسبه می شود. سپس مقادیر

را بررسی می کند، بنابراین رابطه زیر منجر به یک مقدار با ارزش متناظر با تمام ویژگی ها در آن بخش می شود.

$$v_x = \sum_{i=1}^{(2^L)^D} |h_x(P_i)|, \forall x \in W \quad (7)$$

مرتب کردن تمام ویژگی ها در این بخش منجر به یک لیست با اهمیت می شود. اگر $W' \subseteq W$ یک زیرمجموعه انتخاب شده برای شامل شدن ویژگی های با مقادیر متناظر بزرگتر باشد. آنگاه رابطه (۸) می تواند به عنوان یک کلاسیفایر ضعیف برای رد کردن نواحی غیرچهره در دیتابیس آموزشی بکار رود.

$$\theta' = \sum_{x \in W'} h(M(x)) \quad (8)$$

اسکن تصویر برای تشخیص چهره تنها روی نواحی بخش بندی شده پوست با انتخاب بخش های 32×32 پیکسلی انجام می شود. به منظور یافتن تمامی چهره ها با مقیاس های مختلف تغییر ساین تصویر با فاکتور مقیاس ۱.۲ انجام می شود. برای هر پیکسل با توجه به آنالیز همسایه های آن، یک بردار ویژگی بدست می آید. این بردار ویژگی می تواند به صورت یک شاخص محاسبه شود:

$$m = \sum_{i=1}^D V(x_i)(2^L)^{i-1} \quad (9)$$

که در آن $V(x_i)$ میزانی از بردار ویژگی در موقعیت i می باشد. سپس برای تعیین نواحی چهره از کلاسیفایر SNOW با استفاده از جدول جستجو، استفاده می شود [۱۰].

ب- ردیابی چهره

ردیابی چهره شامل دو مرحله می باشد: مرحله اول یادگیری زیرفضای با بعد کاهش یافته مربوط به چهره و به روز رسانی آن با توجه به تغییر در حالات و زاویه چهره و تغییرات نوری در محیط و مرحله دوم استفاده از فیلتر ذره ای براساس یک معیار خطا و تشخیص مرحله ای چهره برای تعیین مکان دقیق و زاویه چهره در فریم های مختلف، برای ردیابی می باشد.

با توجه به اینکه در چهره تشخیص داده شده در طول فریم های مختلف، تغییر حالت و زاویه و مقیاس رخ می دهد، بنابراین مدل الگوی چهره ایجاد شده برای ردیابی توسط فیلتر ذره ای باید با توجه به این تغییرات به روز شود. اگر فرض کنیم مجموعه تصاویر آموزشی چهره $\{I_1, \dots, I_n\}$ باشند، مدل چهره با محاسبه بردارهای ویژه U ماتریس کواریانس $\frac{1}{n-1} \sum_{i=1}^n (I_i - \bar{I})(I_i - \bar{I})^T$ که در آن \bar{I} میانگین نمونه تصاویر آموزشی است، آموزش داده می شود. مدل چهره باید با توجه به منظرهای جدید چهره در هر مرحله تطبیق داده شود. بنابراین اگر فرض کنیم تصاویر آموزشی $A = [I_1, I_2, \dots, I_n]$ و تصاویر آموزشی جدید $B = [I_{n+1}, I_{n+2}, \dots, I_{n+m}]$ و C ترکیب این تصاویر باشد، همچنین میانگین A, B, C با $\bar{I}_A, \bar{I}_B, \bar{I}_C$ نمایش داده شود. برای به روز کردن میانگین و واریانس مطابق با تغییر حالت و جهت چهره و تغییر شرایط نوری به این صورت عمل می شود که U و Σ با استفاده از SVD مربوط به $(A - \bar{I}_A)$ ، همچنین U' و Σ' با

این ویژگی ها نسبت به شکل گیری تصویر $I(x)$ که از انعکاس $R(x)$ و روشنایی $E(x)$ ایجاد می شود، قابل طراحی می باشند. همچنین اثر دوربین می تواند به عنوان یک فاکتور بهره g و ترم بایاس b مدل شود. بنابراین یک مدل تصویر می تواند بوسیله $I(x) = gE(x)R(x) + b$ توصیف شود. به منظور طراحی یک کلاسیفایر مقاوم برای تشخیص باید مولفه انعکاس استخراج شود، زیرا شامل ساختار هدف می باشد. در حالت کلی جداسازی انعکاس و روشنایی مشکل است. یک روند رایج برای حل این مسئله شامل این فرض است که $E(x)$ بصورت فضایی هموار است. اگر فرض شود روشنایی در ناحیه محلی ثابت است $E(x)$ می تواند بصورت رابطه (۲) نشان داده شود.

$$E(x) = E, \forall x \in D \quad (2)$$

ساختار یادگیری SNOW یک شبکه از واحدهای خطی روی فضای ویژگی می باشد. یکی از ویژگی های قوی SNOW توانایی ایجاد یک جدول جستجو برای کلاس بندی می باشد. اگر بخش W از ویژگی های SQMT، $M(x)$ را مورد بررسی قرار دهیم، کلاسیفایر رابطه (۳) با استفاده از جدول غیرچهره $h_x^{nonface}$ و جدول چهره h_x^{face} و

تعریف یک ترشلد برای θ بدست می آید.

$$\theta = \sum_{x \in W} h_x^{nonface}(M(x)) - \sum_{x \in W} h_x^{face}(M(x)) \quad (3)$$

چون هر دو جدول روی یک حوزه کار می کنند، این نشان می دهد که یک جدول جستجوی منفرد میتواند برای کلاس بندی جدول جستجوی منفرد بکار رود:

$$h_x = h_x^{nonface} - h_x^{face} \quad (4)$$

فرض می کنیم دیتابیس آموزشی شامل $i=1, 2, \dots, N$ جزء ویژگی با ویژگی های SQMT باشد و کلاس های مربوط، C_i (چهره و غیرچهره) باشد. جدول چهره و غیرچهره می تواند بوسیله قانون بروز کردن شبکه ای آموزش داده شوند. در ابتدا هر دو جدول شامل صفر می شوند. اگر شاخصی در جدول برای اولین بار در طول آموزش آدرس داده شود، مقدار وزن در آن شاخص به یک تنظیم می شود. سه پارامتر یادگیری: ترشلد γ و پارامتر افزایش رتبه $\alpha > 1$ و کاهش رتبه $0 < \beta < 1$ وجود دارد. اگر $\sum_{x \in W} h_x^{face}(M_i(x)) \leq \gamma$ و C_i یک چهره باشد در این صورت ترفیع رتبه بصورت زیر می باشد:

$$h_x^{face}(M_i(x)) = \alpha h_x^{face}(M_i(x)), \forall x \in W \quad (5)$$

اگر C_i یک غیرچهره باشد و $\sum_{x \in W} h_x^{face}(M_i(x)) > \gamma$ در این صورت کاهش بعد اتفاق می افتد.

$h_x^{face}(M_i(x)) = \beta h_x^{face}(M_i(x)), \forall x \in W$ این روش تا زمانی تکرار می شود که تغییری رخ ندهد. آموزش جدول غیرچهره نیز به همین روش رخ می دهد. در نهایت یک جدول ساده برطبق معادله (۵) ایجاد می شود. برای افزایش سرعت تشخیص از توالی ای از طبقه بندها استفاده می شود.

طبقه بند SNOW تمامی ترکیبات ویژگی ممکن برای یک ویژگی P_i

همان طور که بیان شد برای استفاده از نمایشی که توصیف کننده ویژگی های چهره باشد، تصویر چهره را با استفاده از یک نمایش احتمالی از آنالیز مولفه های اساسی مدل می کنیم. در واقع I_t تصویر چهره پیش بینی شده بوسیله X_t می باشد. فرض می شود I_t از زیرفضایی از هدف که بوسیله U گسترش داده شده و μ مرکز آن است.

احتمال اینکه یک نمونه از این زیر فضا ایجاد شده باشد بطور معکوس با فاصله d یعنی فاصله نمونه تا نقطه مرجع از زیرفضا متناسب است که فاصله تا زیرفضا با d_t و فاصله در داخل زیرفضا (از نمونه تصویر شده تا مرکز زیرفضا) با d_{ω} نمایش داده می شود.

احتمال اینکه یک نمونه از یک زیرفضا ایجاد شده باشد بوسیله یک توزیع گوسی تعیین می شود:

$$p_{d_t}(I_t | X_t) = N(I_t; \mu, UU^T + \mathcal{E}) \quad (14)$$

که در آن I ماتریس واحد و \mathcal{E} مربوط به نویز گوسی در مشاهدات می باشد. در داخل هر زیرفضا، احتمال نمونه تصویر شده به صورت زیر مدل می شود:

$$p_{d_{\omega}}(I_t | X_t) = N(I_t; \mu, U\Sigma^{-2}U^T) \quad (15)$$

که در آن μ مرکز زیرفضا و Σ ماتریس مقادیر ویژه متناظر با ستون های U می باشد. بنابراین احتمال اینکه یک نمونه از زیر فضای خاصی ایجاد شده باشد بصورت زیر می باشد:

$$p(I_t | X_t) = p_{d_t}(I_t | X_t) p_{d_{\omega}}(I_t | X_t) \quad (16)$$

$$= N(I_t; \mu, UU^T + \mathcal{E}) N(I_t; \mu, U\Sigma^{-2}U^T)$$

بنابراین برای ذره ترسیم شده X_t و تصویر متناظر با آن که با I_t نمایش داده می شود، هدف محاسبه $p(I_t | X_t)$ می باشد. برای مینیمم سازی اثر پیکسل های نویزی از تشخیص چهره و نرم خطای $\rho(x, \sigma) = \frac{x^2}{\sigma^2 + x^2}$ استفاده می شود [۱۳]. به این صورت که برای تصحیح در امر پیش بینی و به روز کردن زیرفضای مربوط به چهره و عدم تأثیر برون هشته ها و عواملی که باعث انسداد چهره در امر ردیابی می شوند، از تشخیص چهره بوسیله تشخیص ناحیه به رنگ پوست و همچنین ویژگی های SMQT استفاده می شود. بنابراین پس از ترسیم ذرات مطابق با مدل دینامیکی، برای هر ذره، پنجره مربوط به آن از فریم جاری استخراج شده و احتمال متناظر با مدل مشاهده محاسبه می شود. سپس پنجره متناظر با ذره با بیشترین احتمال ذخیره می شود. هم چنین به روز کردن مدل چهره و میانگین و تشخیص مرحله ای چهره نیز در حین این عمل انجام می شود.

۳- نتایج آزمایشات

مرحله اول برای تشخیص چهره در اولین فریم در فریم های ویدئویی، بخش بندی نواحی به رنگ پوست می باشد. در مرحله بعد، از ویژگی های SMQT و کلاسیفایر SNOW برای تشخیص ناحیه مربوط به چهره تنها در نواحی مربوط به پوست استفاده می شود. اعمال این روش تنها در نواحی ای که به رنگ پوست می باشند باعث افزایش سرعت و کاهش محاسبات به میزان قابل توجهی می شود. به منظور

استفاده از SVD مربوط به $(C - \bar{I}_C)$ محاسبه می شوند. سپس \bar{I}_C و \bar{I}_B با استفاده از رابطه زیر محاسبه می شوند:

$$\begin{aligned} \bar{I}_B &= \frac{1}{m} \sum_{i=n+1}^{n+m} I_i \\ \bar{I}_C &= \frac{n}{n+m} \bar{I}_A + \frac{m}{n+m} \bar{I}_B \end{aligned} \quad (10)$$

با فرض اینکه ابعاد ماتریسهای A و B به ترتیب d^*m و d^*n باشد. ماتریس \hat{B} و \tilde{B} و R به صورت زیر محاسبه می شوند [۱۱]:

$$\hat{B} = [(I_{m+1} - \bar{I}_B) \dots (I_{n+m} - \bar{I}_B)] \sqrt{\frac{nm}{n+m}} (\bar{I}_B - \bar{I}_A) \quad (11)$$

$$\tilde{B} = \text{orth}(\hat{B} - UU^T \hat{B})$$

$$R = \begin{bmatrix} f \Sigma & U^T \hat{B} \\ 0 & \tilde{B} (\hat{B} - UU^T \hat{B}) \end{bmatrix}$$

که در آن $f \in [0, 1]$ فاکتور فراموشی می باشد و به این منظور مورد استفاده قرار می گیرد که تصاویر چهره در فریم های قبلی نسبت به تصاویر چهره در فریم های جدیدتر تأثیر کمتری در به روز کردن میانگین و واریانس داشته باشند. سپس Svd مربوط به R محاسبه می شود که برابر با $\tilde{U} \tilde{\Sigma} \tilde{V}^T$ می باشد و در نهایت \tilde{U} و $U' = [U \ \tilde{B}] \tilde{U}$ محاسبه می شوند.

اگر فرض کنیم متغیر حالت X_t پارامترهای حرکتی هدف را در زمان t توصیف می کند. هدف از ایجاد یک مدل دینامیکی برای فیلتر ذره ای این است که برای تصاویر مشاهده شده $\mathcal{I}_t = \{I_1, \dots, I_t\}$ ، متغیر حالت X_t را تخمین بزنیم. بنابراین رابطه زیر را خواهیم داشت:

$$P(X_t | \mathcal{I}_t) \propto P(I_t | X_t) \int p(X_t | X_{t-1}) p(X_{t-1} | \mathcal{I}_{t-1}) dX_{t-1} \quad (12)$$

بنابراین ردیابی توسط فیلتر ذره ای بوسیله ایجاد مدل مشاهده $P(I_t | X_t)$ و مدل دینامیکی $p(X_t | X_{t-1})$ انجام می شود. برای مدل کردن توزیع روی مکان چهره، از یک الگوریتم بر پایه نمونه گیری براساس ضریب مشاهدات، که یک توزیع اختیاری از مشاهدات را با یک مجموعه ایجاد شده از نمونه های وزن دار تخمین می زند، استفاده می کنیم. برای ایجاد مدل دینامیکی در واقع مکان و موقعیت هدف بوسیله شش پارامتر $X_t = (x_t, y_t, \theta_t, s_t, \alpha_t, \phi_t)$ مشخص می شود. که در آن x_t, y_t انتقال در جهت x و y و θ_t زاویه چرخش، s_t مقیاس، α_t نسبت چهره و ϕ_t جهت خمش مستطیل دربردارنده چهره را نشان می دهد [۱۲]. برای ایجاد مدل دینامیکی در واقع هر پارامتر در X_t بوسیله یک توزیع گوسی حول نقطه متناظر آن در X_{t-1} مدل می شود. بنابراین داریم:

$$p(X_t | X_{t-1}) = N(X_t; X_{t-1}, \Psi) \quad (13)$$

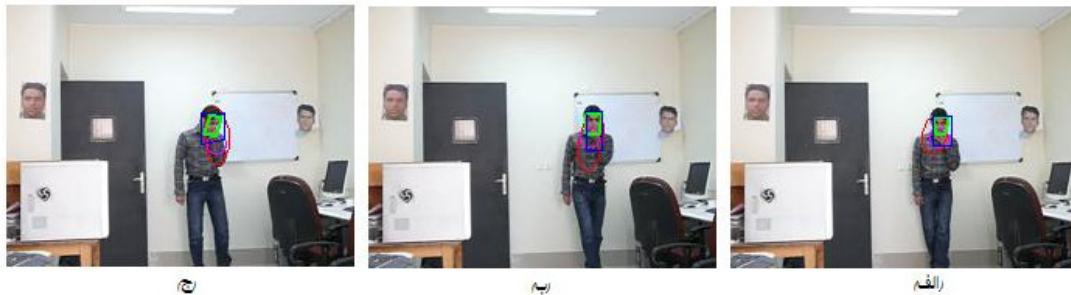
که در آن Ψ یک ماتریس کواریانس قطری است که عناصر آن واریانس های متناظر با پارامترهای مدل می باشند. در واقع این پارامترها نوع حرکت مورد نظر را به ردیاب نشان می دهند. بنابراین باید در فیلتر ذره ای نوعی تعادل بین تعداد ذره های در نظر گرفته شده و کیفیت تخمین توزیع استقرایی وجود داشته باشد.

حالتی که انسداد در چهره وجود دارد، نشان داده شده است. شکل (۴) نیز دلیلی بر پایداری این الگوریتم نسبت به تغییر مقیاس چهره و چرخش سر و تشخیص زاویه چهره و برتری این روش نسبت به سایر روش ها می باشد.

ردیابی چهره در مقابل چهره موجود در عکس و سایر اشیاء، با استفاده از اطلاعات حرکتی (با تفاضل دو فریم) چهره مربوط به شخص تشخیص داده شده و یک مدل مربوط به الگوی چهره ایجاد و با توجه به تغییرات در حالت و زاویه چهره و همچنین تغییر در شرایط نوری به روز می شود. در مرحله بعد، پس از ایجاد مدل دینامیکی و مدل مشاهده، ردیابی به وسیله فیلتر ذره ای با یک معیار خطای مشخص و تشخیص چهره بصورت مرحله ای، انجام می شود. مزیت روش توصیف شده این است که در مقابل حرکات سریع و حتی انسداد چهره بوسیله اشیاء به رنگ پوست مقاوم و با وجود تغییر در شرایط نوری و حالت و مقیاس و زاویه چهره قادر به ردیابی و تعیین زاویه چهره می باشد. بخشی از نتایج آزمایشات در شکل (۳) آمده است. در این شکل نتیجه اعمال الگوریتم پیشنهادی بر روی چند فریم ویدئویی مربوط به دو ویدئوی مختلف برای تشخیص و ردیابی چهره انسان در مقابل چهره موجود در عکس و سایر اشیاء آمده است. همچنین نتایج ردیابی از طریق الگوریتم پیشنهادی با نتایج حاصل از ردیابی بوسیله الگوریتم جابجایی میانگین [۱۴] و جابجایی میانگین بسط یافته براساس هیستوگرام رنگ [۱۵] و فیلتر کالمن [۱۶] مورد مقایسه قرار گرفته است و برتری روش استفاده شده نسبت به این روش ها بخصوص در



شکل (۳): نتایج اعمال الگوریتم ردیابی چهره برای تشخیص و ردیابی چهره در مقابل عکس و سایر اشیاء (مستطیل سبز رنگ) و مقایسه آن با روش جابجایی میانگین (بیضی صورتی رنگ) و روش فیلتر کالمن (مستطیل آبی رنگ) و روش جابجایی میانگین براساس هیستوگرام رنگ (بیضی قرمز رنگ) ردیف اول برای فریم های ۲۸ و ۵۶ و ۱۰۱ و ۱۵۵، ردیف دوم برای فریم های ۵ و ۳۵ و ۸۹ و ۱۲۰.



شکل (۴): نتایج اعمال الگوریتم بر روی فریم های ۱۲ و ۲۵ و مقایسه آن با سایر روشها برای حالتی که در چهره انسداد وجود دارد.

در این مقاله روشی پایدار و بدون درنگ برای تشخیص و ردیابی چهره انسان در مقابل عکس و سایر اشیاء ارائه شد. تشخیص چهره با استفاده از بخش بندی ناحیه چهره و استفاده از ویژگی های SMQT و کلاسیفایر SNOW انجام می شود. برای ردیابی چهره ابتدا چهره انسان از چهره موجود در عکس تشخیص داده شده و سپس ردیابی با اعمال فیلتر ذره ای انجام می شود. در واقع اعمال ویژگی های SMQT تنها بر روی نواحی مربوط به رنگ پوست باعث تسریع در الگوریتم می شود. همچنین به روز کردن مدل چهره با توجه به تغییرات مقیاس و زاویه چهره و تغییر در شرایط نوری باعث مقاوم بودن این روش در مقابل این تغییرات و حتی انسداد در چهره می شود. نتایج آزمایشات دلیل بر پایداری این الگوریتم در مقابل تغییر مقیاس و چرخش سر و با وجود پیش زمینه پیچیده و برتری این روش نسبت به سایر روش ها می باشد.

الگوریتم های جابجایی میانگین و جابجایی میانگین بسط یافته براساس هیستوگرام رنگ قادر به ردیابی دقیق چهره بویژه در هنگام حرکات سریع و وجود انسداد و تغییرات شرایط نوری نیستند. زیرا در این روش ها مدل ایجاد شده برای ردیابی هدف بر اساس هیستوگرام شدت پیکسل ها می باشن. بنابراین هنگام حرکات سریع، انسداد در چهره و عبور چهره از شیء به رنگ پوست مانند عکس در پس زمینه، قادر به تعیین دقیق محدوده چهره نبوده و در مواقعی نیز با انحراف از محدوده مربوط به چهره، در نهایت منجر به ردیابی شیء دیگری به جز چهره می شوند. هم چنین روش مبتنی بر فیلتر کالمن، بخش هایی از گردن فرد را نیز به عنوان چهره در نظر گرفته و ردیابی می کند و در هنگام انسداد نیز دچار مشکل می شود.

۴- نتیجه گیری

مراجع

- [1] S. A. Suandi, S. Enokida and T. Ejima, "Face pose estimation from video sequence using dynamic Bayesian network". IEEE, 2009.
- [2] T. Kabakli, T. Celik, H. Demirel, H. Ozkaramanli, M. Uyguroglu and A.M. Kondo, "Robust face tracking using color segmentation aided with connected component labeling and morphology". From Proceeding (444) Signal and Image Processing - 2004.
- [3] W. Zheng and S. M. Bhandarkar, " Face detection and tracking using a boosted adaptive particle filter", In Proc. Journal of Visual Communication and Image Representation, 2009, pp. 9-27.
- [4] R. C. Verma, C. Schmid, and K. Mikolajczyk, "Face detection and tracking in a video by propagating detection probabilities", In Proc. Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 25. 2003, pp. 1215 –1228.
- [5] Z. Liu, and Y. Wang, "Face detection and tracking in video using dynamic programming", In Proc. International Conference on Image Processing, 2000, pp. 53 – 56.
- [6] M. Vasant, S. Padmanabhan, K. Valentina, B. Matthew, G. Dmitry and K. Rangachar, "A baseline algorithm for face detection and tracking in video". Optics and Photonics for Counterterrorism and Crime Fighting III, Proc. SPIE, Vol. 6741, 2007.
- [7] P. Viola and M. Jones, "Robust real-time face detection", International Journal of Computer Vision, 2004
- [8] J. Yang and C. Liu, "Horizontal and vertical 2D PCA-Based discriminant analysis for face verification on a large-scale database", IEEE Transactions on Information Forensics and Security, Vol 2, 2007.
- [9] E. Sayed and M. Saad, "Frontal-view face detection in the presence of skin-tone regions using a new symmetry approach", JCS&T, Vol. 6, 2006.
- [10] Nilsson, M. Nordberg, J. Claesson, I., " Face Detection using Local SMQT Features and Split up Snow Classifier", IEEE International conference, April 2007.
- [11] R.-S Lin, J. Lim, and M.-H Yang. Adaptive discriminativ generative model and its applications. Advances in Neural Information Processing Systems, Pages 801-808. MIT Press, 2005.
- [12] M. E. Tipping and C.M. Bishop. Probabilistic tracking in metric space. In Proceedings of IEEE International Conference on Computer Vision, Pages 50-57, 2001.
- [13] M.J. Black and A.D. Jepson. "Eigentracking: Robust matching and tracking of articulated objects using view-based representation. In B.Buxton and R.Cipolla, editors, Proceedings of the Fourth European Conference on Computer Vision, LNCS 1064, Springer Verlag, 1996, pp329-342.
- [14] Jianpo Gao, Zhenyang Wu, "Face Tracking Algorithm Based on Mean Shift and Ellipse Fitting", Springer Verlag 2006, pp 270-277.
- [15] R. Collins. "Mean-shift blob tracking through scale space". Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2003.
- [16] Jong-Ho Kim, Byoung-Doo Kang, "Real-Time Face Tracking System Using Adaptive Face Detector and Kalman Filter", Human-Computer Interaction, LNCS 4552, Springer Verlag 2007, pp 669-678.

فاطمه حورعلی در حال حاضر دانشجوی کارشناسی ارشد مهندسی برق (گرایش مخابرات/ سیستم) دانشگاه صنعتی سهند می باشد. زمینه کاری وی پردازش تصویر، بازشناسی الگو، الگوریتم ژنتیک و بینایی کامپیوتر می باشد. آدرس ایمیل: f_hourali@sut.ac.ir

محمدحسین صدیقی مدرک دکتری خود را در سال ۱۹۹۸ از دانشگاه لیورپول دریافت کرد. زمینه های کاری مورد علاقه وی فیلتر کردن مورفولوژی، پردازش امواج ECG، فیلترهای غیرخطی، پردازش سیگنال دیجیتال، پردازش تصویر، پردازش سیگنال M-D، کنترل دیجیتال، الگوریتم ژنتیک، ویولت و فیلتربانک ها، شبکه های عصبی و بازشناسی الگو می باشد. آدرس ایمیل:

sedaaghi@sut.ac.ir