

УДК 57.08

07, июль 2012

Использование дискриминантного анализа и фильтров Габора в задаче распознавания эмоционального состояния по изображениям лица

Латышева Е.Ю.⁽¹⁾, Кашапова Л.Х.⁽²⁾

*Студент⁽¹⁾, аспирант⁽²⁾,
кафедра «Биомедицинские технические системы»*

*Научный руководитель: Спиридонов И.Н., д.т.н., профессор, заведующий кафедрой
«Биомедицинские технические системы»*

МГТУ им. Н.Э. Баумана
kat9-1t@rambler.ru
kashapova.liliya@gmail.com

В современных условиях повышенных психологических нагрузок актуальными задачами являются контроль за психоэмоциональным состоянием и своевременное распознавание психических расстройств у работников профессий, ответственных за жизнь и здоровье людей. Кроме этого, объективная оценка психоэмоционального состояния необходима для психиатрии, где изучение эмоциональной сферы пациента является одним из основных инструментом врача-психиатра.

Одним из наиболее информативных показателей эмоционального состояния являются мимические движения лица. Существует несколько подходов к оценке эмоционального состояния по изображению лица в зависимости от используемого пространства признаков. Пространство признаков, вычисляемое по изображениям лица, может включать в себя геометрические или текстурные признаки [1]. Геометрические признаки представляет собой расстояния, углы и другие особенности, характеризующие особенности расположения и размеров элементов лица. Использование комплекса мимических признаков из 26 расстояний и 7 углов между контрольными точками лица для классификации базовых эмоций позволяет достичь точности распознавания 85 % при правильной расстановке контрольных точек лица [2]. Однако погрешность автоматической расстановки контрольных точек является критическим звеном для анализа видеоизображений [3]. Оператору необходимо контролировать расстановку контрольных точек на каждом кадре видеоизображения и проводить коррекцию в случае необходимости.

В данной статье рассмотрен подход к оценке эмоционального состояния по текстурным признакам изображения лица. Текстуальные признаки изображений вычисляются без привязки к контрольным точкам лица. В качестве текстурных признаков рассмотрены вейвлет-признаки. Наиболее распространенными вейвлетами для анализа изображений лица являются вейлеты Габора, которые

также называют фильтрами Габора [4]. График вейвлета Габора представлен на рис.1.

Банк фильтров Габора представляет собой так называемый «направленный микроскоп». Вейвлет преобразование – это свёртка исходного изображения с некоторой функцией ϕ – вейвлетом. Двумерный вейвлет (ядро, фильтр) Габора, соответствующий ориентации μ и масштабу ν , это:

$$j_{m,n}(x, y) = \frac{\|k_{m,n}\|^2}{S} e^{-\|k_{m,n}\|^2 \|z\|^2 / 2s^2} e^{ik_{m,n}z} - e^{-s^2/2} \Pi \quad (1)$$

где $z = (x, y)$ – координаты вектора, $\|\bullet\|$ – норма вектора, вектор $k_{\mu,\nu} = k_\nu e^{i\theta_\mu}$, $k_\mu = \frac{k_{\max}}{f^\nu}$, $\theta_\mu = \frac{\pi\mu}{8}$, k_{\max} – максимальная частота.

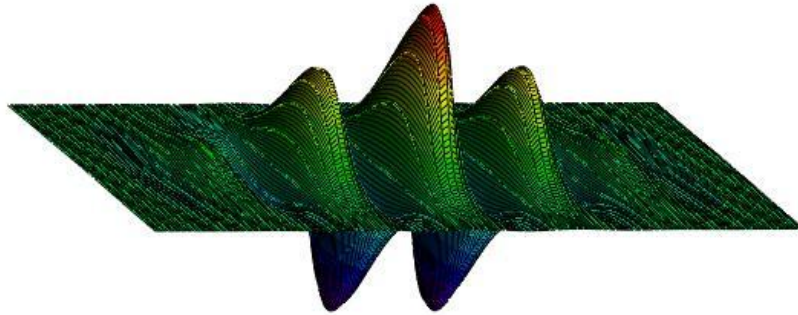


Рис.1. График вейвлета Габора, построенного в соответствии с (1) при $k_{\mu,\nu} = \frac{\pi}{8}$

Обычно используют банк фильтров Габора со следующими характеристиками: 5 различных масштабов $\nu \in \{0, \dots, 4\}$, 8 различных направлений $\mu \in \{0, \dots, 7\}$, $\sigma = \pi$, $k_{\max} = \pi/2$ и $f = \sqrt{2}$ [4].

Изображение $I(x, y)$, подаваемое на вход алгоритма фильтрации, имеет размер $n_r \times n_c$ и является полутоновым. Преобразование Габора изображения представляет собой свёртку изображения с импульсной характеристикой фильтра Габора: $X_{m,n}(x, y) = I(x, y) * j_{m,n}(x, y)$, где $*$ – оператор свёртки.

Пример применения банка фильтров Габора дан на рис.3, где можно наблюдать выделение локальных признаков лица фильтрами в соответствующем направлении и на определённом масштабе. На рис.2 представлено изображение, подаваемое на вход алгоритма.



Рис.1. Пример изображения

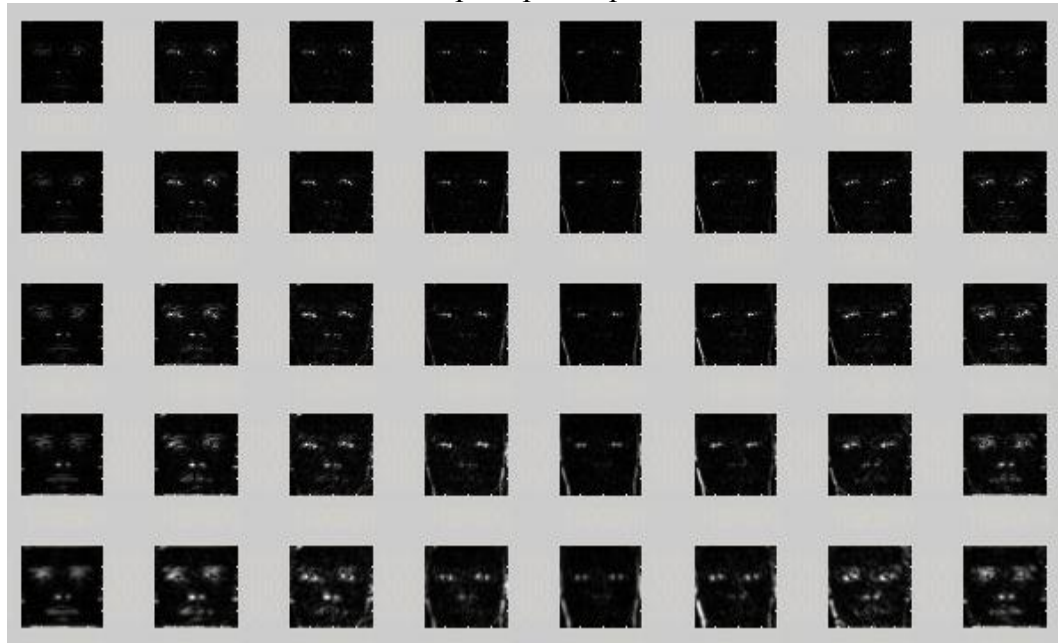


Рис.3. Амплитудная характеристика свёрнутого изображения

После применения банка фильтров Габора для изображений обучающей выборки получаем $N \times O \times S$ матриц признаков, где N – количество изображений обучающей выборки, O – количество ориентаций, S – количество масштабов.

Для того чтобы признаки стали некоррелированными, и матрицы признаков имели меньшую размерность, используется метод главных компонент (2DPCA). Применение 2DPCA к строкам матриц признаков уменьшает количество строк в матрицах. Для перехода в новое пространство признаков необходимо получить ковариационную матрицу[4]: $C = \frac{1}{NOS} \sum_{i,m,n} (X_{m,n}^i - \bar{X})(X_{m,n}^i - \bar{X})^T$, где $\bar{X} = \frac{1}{NOS} \sum_{i,m,n} X_{m,n}^i$ – матрица среднего по всем матрицам признаков обучающей выборки, $()^T$ – операция транспонирования.

Поскольку матрица C размера $n_r \times n_r$ является симметричной, всегда возможно построить ортонормированный базис, состоящий из собственных векторов этой матрицы. При этом будем использовать собственные вектора, соответствующие $d \leq n_r$ ненулевым собственным значениям: $V_d = [\bar{n}_1, \bar{n}_2, \dots, \bar{n}_d]$. Преобразование главных компонент для каждого изображения будет определяться формулой: $Y_{m,n}^i = V_d^T (X_{m,n}^i - \bar{X})$, где $Y_{m,n}^i$ – матрица признаков размера $d \times n_c$, соответствующая μ ориентации и ν масштабу i -го изображения.

Для того, чтобы признаки, соответствующие различным классам, были линейно разделимы, применяем после 2DPCA метод прямого линейного дискриминанта Фишера (2DFLD) к матрицам признаков, уменьшая таким образом размерность матриц по столбцам. Критерий Фишера запишем следующим образом [4]: $J(a) = \frac{a^T S_B^P a}{a^T S_W^P a}$, где S_B^P - матрица межклассового разброса признаков, S_W^P - матрица внутриклассового разброса признаков. Матрицы $S_B^P = \frac{1}{NOS} \sum_{i=1}^K N_i S_O(\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y})^T$, $S_W^P = \frac{1}{NOS} \sum_{i,j,m,n} (Y_{m,n}^{ij} - \bar{Y}_i)^T (Y_{m,n}^{ij} - \bar{Y}_i)$ имеют размерность $n_c \times n_c$, где $\bar{Y}_i = \frac{1}{N_i OS} \sum_{j,m,n} Y_{m,n}^{ij}$ - матрица среднего по матрицам признаков i -го класса, $\bar{Y} = \frac{1}{NOS} \sum_{i,j,m,n} Y_{m,n}^{ij}$ - матрица среднего по матрицам признаков всех классов, K – количество классов.

Определим собственные вектора β и собственные значения \sum матрицы S_B^P : $S_B^P \beta = \sum \beta$, выбираем g собственных векторов, соответствующих g из n_c ненулевым собственным числам. Затем проецируем матрицу S_W^P в подпространство, соответствующее g выбранным собственным векторам: $\beta_g^T S_W^P \beta_g = S_W^B$. Получаем матрицу S_W^B размера $g \times g$, которая может быть диагонализирована: $\gamma^T S_W^B \gamma = \eta$, где γ и η матрицы собственных векторов и собственных значений матрицы S_W^P . Отобрав q из g собственных векторов, формируем итоговую матрицу признаков, соответствующую μ ориентации и ν масштабу i -го изображения: $Z_{m,n}^i = Y_{m,n}^i \beta_g \gamma_q$.

Матрица $Z_{m,n}^i$ имеет размер $d \times q$ и содержит признаки, максимально разделённые в столбцах по классам. Объединяя матрицы признаков $Z_{m,n}^i$, относящиеся к одному изображению i , сформируем комплекс признаков i -го изображения.

К полученному комплексу признаков i -го изображения применим простейший классификатор, вычисляя расстояние между двумя матрицами признаков, где признаки линейно разделены в столбцах. Матрицы признаков, соответствующие двум различным изображениям, с фиксированной ориентацией и масштабом: $Z_{m,n}^n = [Z_{m,n,1}^n, Z_{m,n,2}^n, \dots, Z_{m,n,q}^n]^T$ и $Z_{m,n}^m = [Z_{m,n,1}^m, Z_{m,n,2}^m, \dots, Z_{m,n,q}^m]^T$, расстояние между двумя комплексами признаков, относящихся к двум различным исходным изображениям: $dis(Z^n, Z^m) = \sum_{m,n,k} \|Z_{m,n,k}^n - Z_{m,n,k}^m\|_2$, где $\|Z_{m,n,k}^n - Z_{m,n,k}^m\|_2$ - евклидово расстояние между двумя соответствующими вектор-столбцами матриц признаков $Z_{m,n,k}^n$ и $Z_{m,n,k}^m$. Вычисляя расстояние между комплексом признаков тестируемого изображения, Z^{test} , и N числом комплекса признаков обучающей выборки, Z^w , определяем принадлежность тестируемого изображения классу K , $Z^w \in K$:

$$dis(Z^{test}, Z^w) = \min_j dis(Z^{test}, Z^j)$$

Для формирования обучающей и тестируемой выборки использовалась база JAFFE, состоящая из 213 изображений 10 японских женщин, показывающих 7 эмоций. 7 эмоций представляют 7 классов:

- 0 (neutral/покой) – 30 изображений,
- 1(sad/печаль) – 31 изображение,

- 2(happy/счастье) – 31 изображение,
- 3(fear/страх) – 32 изображения,
- 4(disgust/отвращение) – 29 изображений,
- 5(anger/гнев) – 30 изображений,
- 6(surprise/удивление) – 30 изображений.

Имеется от 2 до 4 изображений женщин каждого класса. Размер каждого изображения 256x256 пикселей, изображения представлены в формате TIFF. Изображения, подаваемые на вход алгоритма фильтрации банком фильтров Габора, подвергаются дополнительной обработке: на каждом изображении выделяется и вырезается область лица, изображение конвертируется в полутоновый формат. В эксперименте обучающая выборка представляла собой 7 классов по 15 изображений в каждом, набранных произвольным образом из базы. Остальные 105 изображений составили тестируемую выборку.

Результаты эксперимента приведены в таблицах 1-3.

Таблица 1

Результаты тестирования: $p = g = q = 20$,

Σ_i - количество изображений i -го класса тестируемой выборки, % - процент распознавания

	0	1	2	3	4	5	6	Σ_i	%
0	11			1			3	15	73,3
1	1	10		2				13	76,9
2	3	4	9					16	56,3
3	1	2	1	9	2		2	17	52,9
4				1	10	2	1	14	71,4
5	3	2			1	9		15	60,0
6				3	1		11	15	73,3

Таблица 2

Результаты тестирования: $p = g = q = 10$,

Σ_i - количество изображений i -го класса тестируемой выборки, % - процент распознавания

	0	1	2	3	4	5	6	Σ_i	%
0	11						4	15	73,3
1	1	10		2				13	76,9
2	3	2	11					16	68,7
3	1	3		9	2		2	17	52,9
4					12	1	1	14	85,7

5	3	1			1	10		1	66,71
6				2			12	15	80,0

Таблица 3

Результаты тестирования: $p = g = q = 5$,

Σ_i - количество изображений i -го класса тестируемой выборки, % - процент

распознавания								Σ_i	%
	0	1	2	3	4	5	6		
0	12						2	1	80,0
1	1	11		1				13	84,6
2	2	2	12					16	75,0
3	1	3		9	2		2	17	52,9
4		2	1		11			14	78,6
5	4	1			1	9		15	66,71
6	2			2			11	15	80,0

При рассмотрении таблиц результатов можно сделать вывод, что наилучший процент распознавания эмоций печали, счастья и нейтрального состояния достигается при $p = g = q = 5$, эмоции отвращения при $p = g = q = 10$, эмоций гнева, удивления при $p = g = q = 5$ и $p = g = q = 10$.

Результаты показывают, что фильтры Габора могут быть использованы для оценки психоэмоционального состояния по изображению лица. Вероятность правильного распознавания на тестовой выборке составила 72,38%. По сравнению с алгоритмом, основанным на расстановке контрольных точек лица, алгоритм с использованием фильтров Габора не требует контроля и коррекции со стороны оператора.

Список литературы

1. Зинин А.М. Руководство по портретной экспертизе: учебное пособие. – М.: Эксмо, 2006. – 208 с.
2. Кашапова Л.Х., Крылов Ю.В., Спиридонов И.Н. Разработка алгоритма оценки эмоционального состояния по видеоизображению лица // IX МНТК «Физика и радиоэлектроника в медицине и экологии» ФРЭМЭ'2010: Материалы НМШ по биометрическим технологиям для медико-биологических анализов и функциональной диагностики. – Владимир, 2010. – С. 83-88.
3. Кашапова Л.Х., Хрулев А.А., Спиридонов И.Н. Использование видеоизображений для автоматизированной оценки мимических изменений // Медико-технические технологии на страже здоровья: Сб. докладов X РНТК, г.

Монастир (Тунис), 28 сентября – 05 октября 2008 г. – М.: НИИ РЛ МГТУ им. Н.Э. Баумана, 2008. – С.76 – 78.

4. R.M. Mutelo, W.L. Woo, S.S. Dlay Discriminant analysis of the two-dimensional Gabor Features for face recognition//The Institution of Engineering and Technology. – 2008. – С.37 – 49.