

Использование спектрального преобразования для распознавания напечатанного изображения*

Пушняков А. С.

aleksey.pushnyakov@phystech.edu

Московский физико-технический институт

Аннотация. В данной работе решается задача классификации двух типов изображений глаз: реально сфотографированного и впоследствии распечатанного. Используется метод спектрального преобразования изображения. В напечатанном изображении предполагается обнаружить периодическую структуру, которая порождает дополнительную гармонику в спектре. Рассматривается радиальная составляющая фурье-спектра, и по ней строится пространство признаков. Задача классификации решается с помощью метрического классификатора.

Ключевые слова: *iris recognition, дискретное преобразование Фурье, метрический классификатор.*

Recognition of printed image using spectral transform*

Pushnyakov A. S.

Moscow Institute of Physics and Technology

This paper is about a problem of classification real and printed eye-image. A spectral transform of image is used. We proposed to find a periodical structure that generates high frequency spectral magnitude. A feature space is based on radial component of the Fourier image. The problem of classification is solved by metric classifier.

Keywords: *iris recognition, discrete Fourier transform, metric classifier.*

Введение

Задача классификации напечатанных и реальных изображений возникает в более общей задаче распознавания человека по радужной оболочке глаза [1, 2]. Изображение человеческого глаза может быть подделано. Одним из возможных способов такой фальсификации является фотографирование глаза и последующая печать изображения. Возникает вопрос, по каким критериям можно отличать два типа изображений при непосредственном сканировании. Предполагается, что при печати изображения появляется периодический шум высокой частоты (рис. 1). Данную частоту можно выделить с помощью спектрального преобразования изображения.

Спектральные преобразования используются в различных типах фильтрации изображений, при сжатии изображений с минимальными потерями качества и в других аспектах обработки изображений [4, ?]. В нашем случае предлагается использовать дискретное преобразование Фурье [3, 5]. Одним из примеров использования преобразования Фурье является задача распознавания изображений с учетом инвариантности относительно поворотов и масштабирования [6, 7].

Проблема фальсификации изображений при распознавании человека по радужке рассматривается в [8]. В числе прочего в этой работе поднимается проблема классификации

Научные руководители И.А. Матвеев, В.В. Стрижов

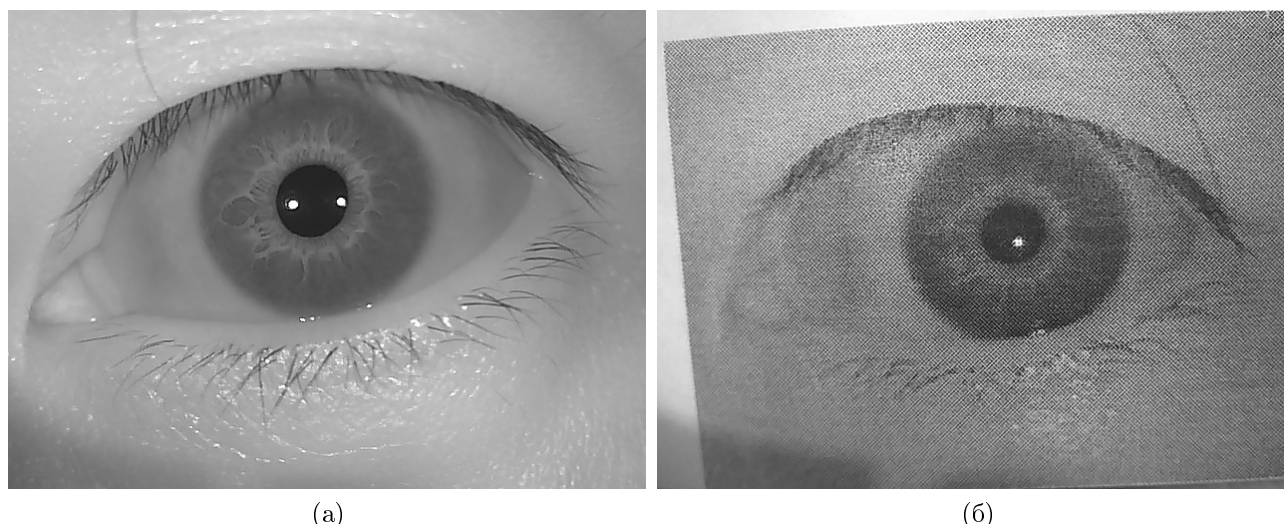


Рис. 1: (а) сфотографированное изображение глаза; (б) напечатанное изображение глаза

напечатанного и реального изображения. Некоторые решения, основанные на спектральном анализе, предложены в работах [9, 10].

В нашем случае коэффициенты Фурье составляют пространство первичных признаков. Используя эти признаки, предлагается классифицировать два выше описанных типа изображений. Так как размерность пространства первичных признаков велика, предлагается построить новые признаки, которые уже будут непосредственно использоваться для классификации. Для построения новых (вторичных) признаков используется радиальная составляющая фурье-образа. Во-первых, такой подход может на порядок сократить размерность задачи классификации. Во-вторых, в силу того что предполагаемая выделяющаяся гармоника имеет приблизительно равные пространственные частоты, то при переходе к радиальной компоненте эти гармоники должны наложиться друг на друга. В итоге задача сводится к задаче классификации, которую, например, можно решить с помощью метрического классификатора [11], в частности, используя метод Парзенковского окна [12].

В работе также проводится оценка сложности предлагаемого алгоритма и проводится вычислительный эксперимент, проверяющий быстроедействие и качество классификатора.

Постановка задачи

Рассмотрим наше изображение как сеточную функцию $f(x, y)$, где $x = 0, 1, \dots, M - 1$, $y = 0, 1, \dots, N - 1$. Дискретное преобразование Фурье можно записать в следующем виде

$$\varphi(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi i \left(\frac{ux}{M} + \frac{vy}{N} \right)} \quad (1)$$

Элементы матрицы $\Phi = \|\varphi(u, v)\|_{u=0, v=0}^{M-1, N-1}$ — суть пространство первичных признаков. Под вторичным признаком понимается некоторая функция $g(\Phi)$. Задача состоит в том, чтобы подобрать достаточно небольшой набор вторичных признаков $\mathbf{g} = \{g_k\}_{k=1}^I$, которые бы отражали распределение плотности спектра Φ (при этом предполагается, что I значительно меньше, чем MN). В частности, нас интересует наличие максимумов в области высоких частот. Наличие таких максимумов говорит о наличии периодического шума высокой частоты, который наблюдается в поддельных напечатанных изображениях.

Для классификации изображений используется метрический классификатор следующего вида

$$a(u, X_+, X_-) = \text{sign}(W(u, X_+) - W(u, X_-)), \quad (2)$$

где u — классифицируемое изображение, X_+ и X_- — обучающие выборки настоящих и поддельных изображения соответственно, функция $W(u, X)$ определяет принадлежность объекта u к классу X .

Описание алгоритма

Пусть наше изображение $f(x, y)$ преобразовано согласно (1) в $\varphi(u, v)$ с помощью быстрого преобразования Фурье [13].

Первым шагом введем понятие радиальной компоненты фурье-образа Φ . Отметим, что нулевая гармоника спектра в нашем представлении есть $\varphi(0, 0)$ (малые частоты обычно самые интенсивные). Поэтому точку $(0, 0)$ примем за полюс. Далее сгруппируем все точки (u, v) по евклидовому расстояния до $(0, 0)$ с учетом тороидального замыкания сетки. Обозначим $\hat{u} = \min\{u, M - 1 - u\}$, $\hat{v} = \min\{v, N - 1 - v\}$ и рассмотрим множество $S(r) = \{(u, v) : r^2 \leq \hat{u}^2 + \hat{v}^2 < (r + 1)^2\}$. Здесь также учтено, что мы хотим иметь целые значения r . Тогда можно ввести радиальную компоненту фурье-образа как

$$R(r) = \frac{1}{|S(r)|} \sum_{(u,v) \in S(r)} |\varphi(u, v)| \quad (3)$$

При этом $r \leq r_{\max} = \frac{1}{2}\sqrt{M^2 + N^2}$.

Теперь опишем процедуру получения вторичных признаков $\{g_k\}_{k=1}^I$. Для этого введем понятие интегрального признака для радиальной компоненты фурье-образа. Интегральным признаком $\theta(\alpha)$, $0 < \alpha < 1$ для $R(r)$ назовем следующую величину

$$\theta(\alpha) = \max \left\{ \theta : \sum_{r=0}^{\theta} R(r) \leq \alpha \sum_{r=0}^{r_{\max}} R(r) \right\} \quad (4)$$

Фактически, интегральный признак показывает то значение аргумента r , при котором площадь подграфика спектра есть доля α от общей площади подграфика. Основная идея состоит в том, что данные интегральные признаки отображают распределение плотности спектра. Поэтому для спектров настоящих изображений, которые почти монотонно убывают, и спектров поддельных изображений, которые имеют четко выделенные пики высокой частоты, эти характеристики должны значительно отличаться. Однако априори не понятно, какие значения α нужно выбирать. Также нужно учесть, что значительная часть плотности сосредоточена вблизи нулевых частот.

Предлагается построить последовательность α_k такую, чтобы последовательность $\delta_k = \alpha_k - \alpha_{k-1}$ достаточно быстро убывала. Положим $\delta_k = 2^{-k}$, $\alpha_0 = 0$ и соответственно $\alpha_k = \sum_{j=1}^k 2^{-j} = 1 - 2^{-k}$ (вообще говоря, можно брать δ_k как любую бесконечно убывающую геометрическую прогрессию). В силу дискретности задачи существует такой номер I , что $\theta(\alpha_I) = r_{\max}$ и на этом последовательность обрывается. Тогда определим $g_k = \theta_k - \theta_{k-1}$, $k = 1, \dots, I$. Таким образом, вторичные признаки g_k есть скачки интегральных признаков. А именно: при изменении аргумента с θ_k на величину g_k значение площади подграфика увеличивается на долю δ_k от общей площади. Большое значение g_k указывает на наличие пика в спектре. Отметим, что, вообще говоря, I зависит от исследуемого

изображения, но последовательность θ_k можно продолжить стационарным образом, поэтому достаточно для всех изображений установить единый порог I_* . Например, его можно выбрать как максимальное значение I на обучающей выборке.

Также отметим, что можно каждый раз не пересчитывать сумму $\sum_{r=0}^{\theta} R(r)$, а определять $\theta_k = \theta(\alpha_k)$ следующим образом

$$\theta_k = \max \left\{ \theta : \sum_{r=\theta_{k-1}}^{\theta} R(r) \leq \frac{1}{2} \sum_{r=\theta_{k-1}}^{r_{max}} R(r) \right\} \quad (5)$$

В качестве метрики на наборах $\mathbf{g}$ выбирается евклидова метрика в $\mathbb{R}^I$

$$\rho(\mathbf{g}_1, \mathbf{g}_2) = \left(\sum_{i=0}^I |g_{1i} - g_{2i}|^2 \right)^{\frac{1}{2}}, \quad (6)$$

Для классификации используется метод Парзенковского окна с экспоненциальным ядром $K(x) = e^{-x^2}$. Ширина окна h выбирается равной максимальному расстоянию между классифицируемым объектом u и элементом обучающей выборки $x \in X_+ \cup X_-$

$$h(u) = \max_{x \in X_+ \cup X_-} \hat{\rho}(u, x) \quad (7)$$

где $\hat{\rho}(u, x) = \rho(\mathbf{g}(u), \mathbf{g}(x))$ определяется из (6). Отсюда получаем следующую функцию принадлежности объекта u классу X

$$W(u, X) = \frac{1}{|X|} \sum_{x \in X} K \left(\frac{\hat{\rho}(u, x)}{h(u)} \right), \quad (8)$$

и классификатор принимает следующий вид

$$a(u, X_+, X_-) = \text{sign} \left(\frac{1}{|X_+|} \sum_{x \in X_+} K \left(\frac{\hat{\rho}(u, x)}{h(u)} \right) - \frac{1}{|X_-|} \sum_{x \in X_-} K \left(\frac{\hat{\rho}(u, x)}{h(u)} \right) \right) \quad (9)$$

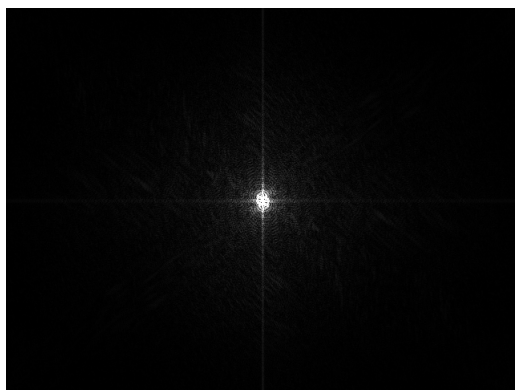
Оценка сложности алгоритма

Покажем что, время работы как классификатора, так и обучающего алгоритма определяется временем работы спектрального преобразования. Сложность быстрого преобразования Фурье есть $O(L \log L)$, где $L = NM$ — количество пикселей в изображении. В это время сложность остальной части алгоритма является линейной по L . Действительно, вычисление радиальной компоненты согласно формуле (3) выполняется за не более чем $3MN$ арифметических операций (с учетом разбиения всех точек на классы $S(r)$). Построение каждого интегрального признака θ_k и вместе с ним g_k не более чем за $2r_{max} = \sqrt{M^2 + N^2}$ арифметических операций (радиальная компонента $R(r) = 0$, при $r > r_{max}$). При этом количество признаков I фиксировано и мало по сравнению с L . Сложность самой классификации определяется только размером обучающей выборки l и количеством признаков I (также l значительно меньше L). Таким образом, общую сложность алгоритма можно оценить сверху как $O(L \log L)$.

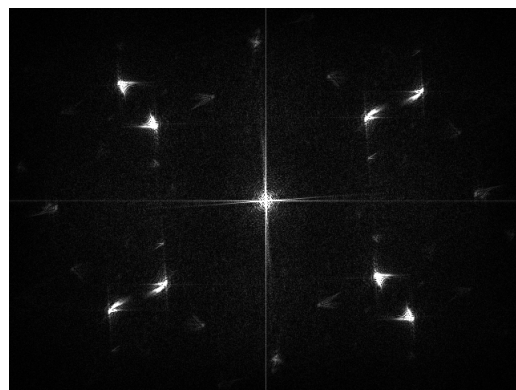
Вычислительный эксперимент

В вычислительном эксперименте использовались изображения в формате jpeg с разрешением 640×480 , то есть в формуле (1) нужно положить $N = 640$, $M = 480$. Изображения были разделены на 4 группы изображений: две группы использовались как обучающая выборка, две другие — контрольная выборка. В обучающей выборке было использовано 1000 реальных и 1000 поддельных изображений, в двух контрольных выборках было соответственно по 3000 изображений. Также классификация была проведена на меньших размерах выборок (порядка 100 изображений)

Как и ожидалось, наличие периодического шума приводит к сильному различию спектров полученных после дискретного преобразования Фурье согласно формуле (1). На рис. 2 изображен результат преобразования Фурье для изображения глаз на рис. 1.



(а) реальное изображение



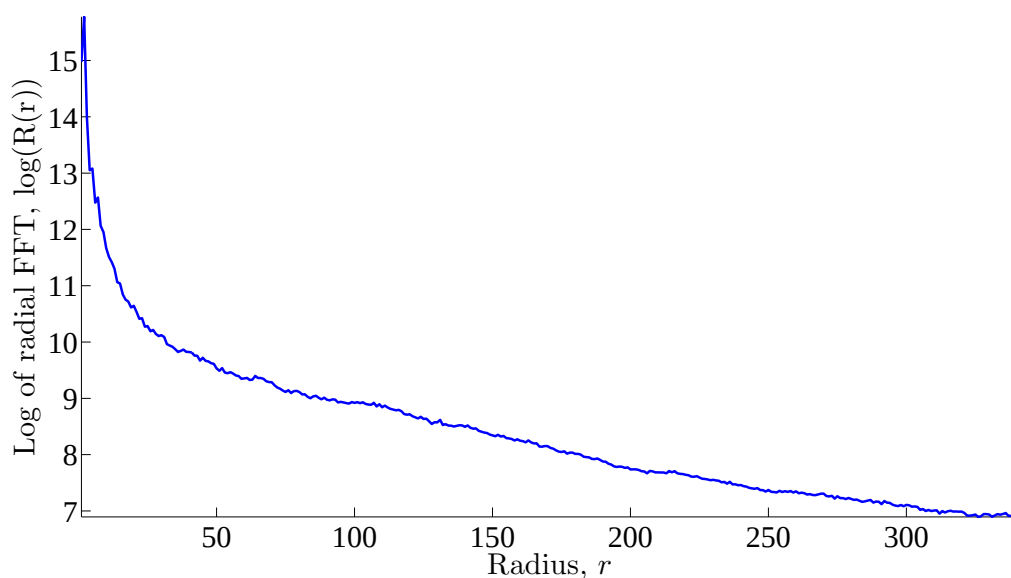
(б) напечатанное изображение

Рис. 2: Двумерный фурье-образ изображения

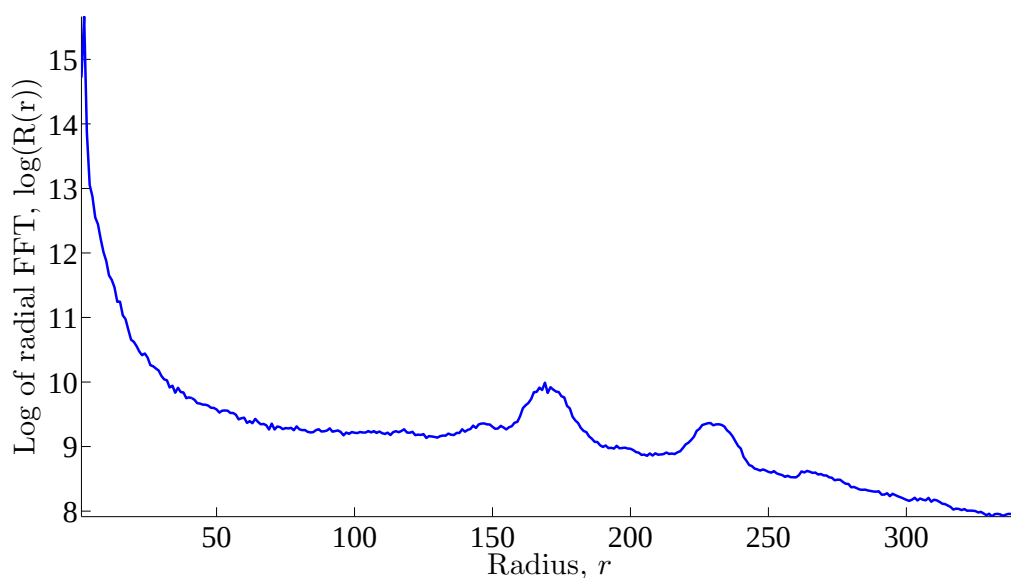
Как видно из последнего рисунка, в спектре напечатанного изображения есть 8 побочных максимумов; причем есть две группы по 4 максимума, находящиеся на приблизительно одинаковом расстоянии от нулевой частоты. При построении радиальной компоненты фурье-образа по формуле (3) мы получаем два пика у спектра напечатанного изображения. Для приведенных выше спектров радиальные компоненты приведены в логарифмическом масштабе на рис. 3.

Как показал эксперимент, для выявления различий достаточно выбрать очень небольшое количество признаков. Во-первых, почти для всех изображений значение I (номер интегрального признака, после которого последовательность θ_k стабилизируется) не превышал 10. Во-вторых, значительные различия во вторичных признаках g_k возникают уже при $k < 5$. Поэтому в эксперименте полагалось $I = 10$. График значений g_k для тестовой обучающей выборки из 20 изображений показан на рис. 4. Каждая ломанная соответствует набору признаков одного изображения. Резкий пик для напечатанных изображений (красный цвет) при $k = 3 \div 4$ указывает на наличие побочного максимума в спектре.

Экспериментально была проверена скорость работы алгоритма. При этом оказалось, что она почти не зависит от размера обучающей выборки. Так в основном эксперименте (обучающая выборка 2000 изображений) время в секундах работы классификатора в основном эксперименте (6000 изображений) составило $T_1 = 1704$, т.е. время обработки одного изображения $t_1 = 0,284$, а для обучающей выборки из 200 изображений $t_2 = 0,235$.



(а) реальное изображение



(б) напечатанное изображение

Рис. 3: Радиальная компонента фурье-образа

Время работы обучающего алгоритма на количество элементов выборки составило $\tau_1 = 0,241$, $\tau_2 = 0,225$ для выборок размера 4000 и 200 соответственно. Зависимость времени работы классификатора (на одном изображении) от размера длины обучающей выборки l показана на рис. 5. Из рисунка видно, что время работы вначале растет, а потом флуктуирует около значения $\hat{t} = 0,265$. Эти результаты согласуются с тем, что время работы алгоритма определяется сложностью спектрального преобразования и почти не зависит от размеров обучающей выборки.

Основной причиной возникновения ошибок при работе классификатора являлось отсутствие фокусировки и значительного числа изображений. Это связано с тем, что изображения получались сериями (съемка камерой), среди которых лишь одно или два изоб-

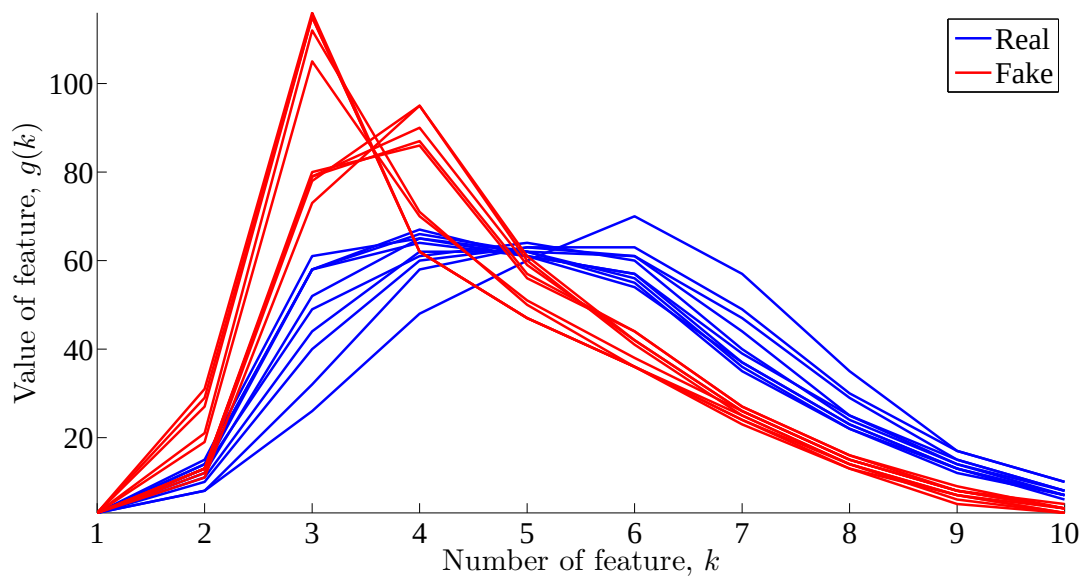


Рис. 4: Значения признаков g_k на тестовой обучающей выборке

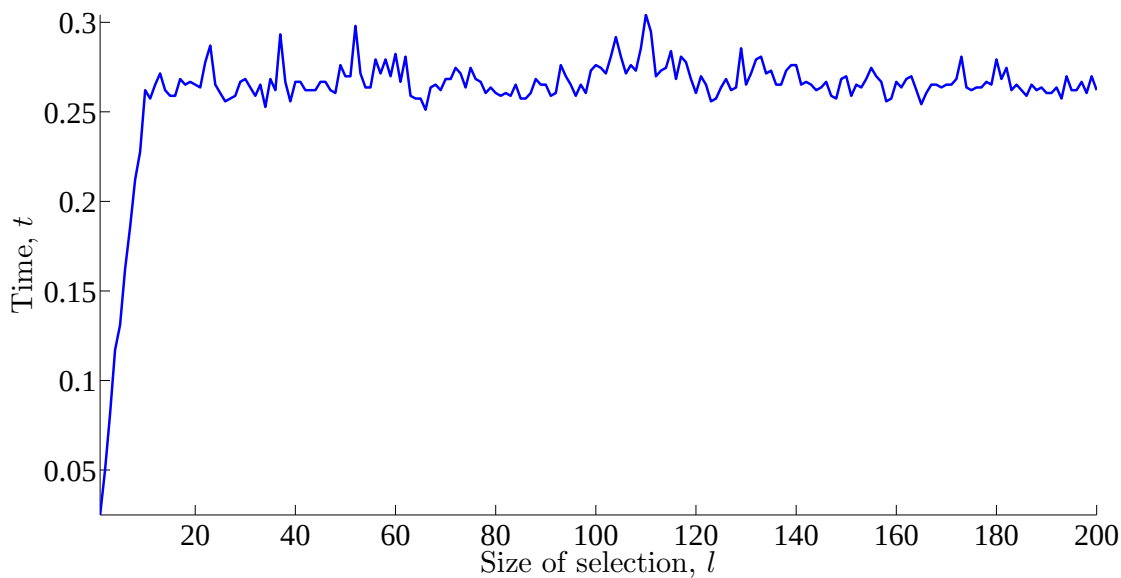


Рис. 5: Зависимость времени работы классификатора от размера обучающей выборки

ражения попадают в фокус. На хорошо сфокусированных фотографиях и при большой обучающей выборке (более 100 изображений) доля ошибок не превосходила 4–6% (это относится к ошибочным классификациям как поддельных, так и настоящих изображений). В общем случае при длине обучающей выборки более 100 изображений и наличии в ней плохо сфокусированных изображений доля ошибок не превосходила 10–15%. Однако на малых выборках (менее 50 изображений) доля ошибок второго рода (поддельное изображение распознается как настоящее) доходила до 25–30%. Возможными решения данной проблемы является предварительная фильтрация изображений, а также обработка изображений одного глаза сериями. В последнем случае серия определяется как напечатанная, если хотя бы на одно из изображений в серии определено как напечатанное.

Заключение

Предложен алгоритм классификации напечатанных и реальных изображений глаз, использующий дискретное преобразование Фурье. В качестве признаков для квалификации используются интегральные характеристики радиальной компоненты фурье-образа. Проведен вычислительный эксперимент, показывающий быстроедействие алгоритма, достаточное для реального применения. Для повышения эффективности и надежности работы алгоритма возможна оптимизация метрического классификатора, например, с помощью выделения эталонных объектов в обучающей выборке.

Литература

- [1] *Daugman J.* How Iris Recognition Works // *IEEE Transactions on Circuits and Systems for Video Technology*. 2002. Vol. 14. Pp. 21–30.
- [2] *Wildes R.P.* Iris recognition: an emerging biometric technology // *Proceedings of the IEEE*. Vol. 85, № 9. Pp. 1348–1363
- [3] *Гонсалес, Р. and Вудс, Р.* Мир цифровой обработки Цифровая обработка изображений. ТЕХНО-СФЕРА. 2005.
- [4] *Сорока Е. З., Хлбородов В. А. and Хуанг Т.* Обработка изображений и цифровая фильтрация. Мир. 1979.
- [5] *Форсайт Д., Понс Ж.* Компьютерное зрение. Современный подход. ИД Вильямс М. 2004.
- [6] *Hsu et Al.* Rotation-invariant digital pattern recognition using circular harmonic expansion // *Applied Optics*. 1982. Vol. 21, № 22. Pp. 4012–4015.
- [7] *Reddy B. S., Chatterji B. N.* An FFT-based technique for translation, rotation, and scale-invariant image registration // *Image Processing, IEEE Transactions on*. 1996. Vol. 5, № 8. Pp. 1266–1271.
- [8] *Daugman J.* Recognising persons by their iris patterns // *Advances in Biometric Person Authentication*. 2005. Pp. 783–814.
- [9] *He et Al.* A fake iris detection method based on fft and quality assessment // *Pattern Recognition, CCPR'08. Chinese Conference on*. 2008. Pp. 1–4.
- [10] *He et Al.*, A new fake iris detection method // *Advances in Biometrics*. 2009. Pp. 1132–1139.
- [11] *Ту, Дж. и др.* Принципы распознавания образов: Пер. с англ. Мир. 1978.
- [12] *Pascal V., Yoshua B.* Manifold parzen windows // *Advances in Neural Information Processing Systems*. 2002. Vol. 15. Pp. 825–832.
- [13] *Johnson, Steven G., Frigo M.* A modified split-radix FFT with fewer arithmetic operations // *Signal Processing, IEEE Transactions on*. 2007. Vol. 55, № 1. Pp. 111–119.