

Использование фильтра Габора и гиперболического тангенса в нейросети выделения краев*

А. Кугаевских

a-kugaevskikh@yandex.ru

Россия, Тюменский государственный университет

В работе приведена структура слоя выделения границ для нейросетевой модели сегментации изображений. Представлены 3 типа рецептивных полей нейронов, базирующихся на использовании фильтра Габора и гиперболического тангенса. Проведено сравнение различных коэффициентов и параметров и выявлена оптимальная формула ядра фильтра Габора и оптимальные значения параметров.

Ключевые слова: сегментация изображений, выделение границ, нейронная сеть, фильтр Габора

1. Введение

Построение нейронных сетей для задач компьютерного зрения предполагает использование большой базы признаков на первых слоях, выделенных перед этим сегментацией изображения. Эти этапы можно совместить в одной нейронной сети. Хорошей попыткой воплощения этой идеи является архитектура HMAX [1]. Недостатком этой модели можно считать использование для выделения краев только фильтра Габора, ориентированного больше на выделение темных полосок в окружении светлого фона. Помимо HMAX, задачу нейросетового выделения краев пытаются решить с помощью многослойного персептрона [2, 3], самоорганизующихся карт Кохонена [4] и PCNN [5].

2. Структура слоя

Первый слой детектора краев предлагается строить из нейронов двух типов рецептивных полей. Нейроны одного типа с поворотом рецептивного поля на каждые 10 градусов сгруппированы в ориентационные колонки. Ориентационные колонки всех типов, в свою очередь, группируются в суперколонку. В свою очередь, каждая суперколонка обрабатывает определенную область изображения.

Фактически нейроны реализуют операцию свертки изображения с ядром фильтра. Ядро фильтра задает весовые коэффициенты связей нейрона с соответствующими пикселями изображения. Пороговая функция такого нейрона позволяет снизить шум.

$$u = \max(\nu, 0),$$

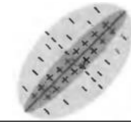

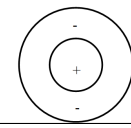
$$\nu = \sum_{|k,l| < A_g} u_0(m+k, n+l)G_t, \quad (1)$$

где G_t – матрица весов рецептивного поля типа t , m, n – позиция центра рецептивного поля, k, l – позиция светового импульса в рецептивной области.

Внутри ориентационной колонки нейроны функционируют по принципу «победитель получает все»

(WTA). В совокупности с функцией активации это позволяет избежать необходимости использования тормозных связей внутри колонок и суперколонок.

Таблица 1: Типы рецептивных полей.

№	Рисунок	Функция	Параметры
1		Фильтр Габора	$\gamma = 0.1$
2		Гиперболический тангенс $G_2 = th(\frac{2x'}{\lambda})$	
3		Фильтр Габора	$\gamma = 1$

3. Фильтр габора

Ядро фильтра Габора является произведением гауссиана и гармонической функции, и фильтрует сигнал, основываясь на параметре предпочтительной пространственной частоты λ^{-1} :

$$G_1 = \exp(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}) \cos(2\pi \frac{x'}{\lambda}) \quad (2)$$

Эта формула [6] имеет оптимальную постоянную компоненту фильтра. Проанализировав многие работы по применению фильтра Габора можно составить обобщенную формулу:

$$G_1 = C_1 \exp(-C_2 \frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}) \cos(C_3 \frac{x'}{\lambda}) \quad (3)$$

Масштабирующие коэффициенты определяют форму кривой ядра фильтра и у разных авторов принимают разные значения, которые представлены в таблице 2.

Работа опубликована по гранту РФФИ №16-07-20482.

C_1 непосредственно определяет селективную способность фильтра. С ростом λ ядра с параметрами 1-4 вырождаются практически в прямую, что негативным образом сказывается на селективной способности. Ядра с параметрами 5-6 вырождаются гораздо медленнее.

Уменьшение λ приводит к вырождению ядер с параметрами 3-5, ядро с параметром 6 при этом дает слишком большой экстремум.

Таблица 2: Коэффициенты обобщенной формулы.

C_1	C_2	C_3
$\frac{\lambda^{-2}}{\sigma_x \sigma_y}$	λ^{-2}	-2π
$\frac{\lambda^{-2}}{\pi \sigma_x \sigma_y}$	$\frac{1}{2}$	2π
$\frac{\lambda^{-2}}{2\pi \sigma_x \sigma_y}$	1	1
$\frac{\lambda^{-2}}{\sqrt{2\pi \sigma_x \sigma_y}}$	π	-1
1	$\frac{\lambda^{-2}}{2}$	
$\frac{\sqrt{2}}{\sigma}$		

C_2 определяет форму гауссиана. При $C_1 = 1$ с ростом λ , ядра с параметрами 1 и 5 практически совпадают, ядро с параметром 4 вырождается в прямую, ядро с параметром 2 становится ближе к 1 и 5. С ростом λ , и соответственным изменением σ_x, σ_y , уменьшается отрицательная зона у ядер с параметрами 1, 4, и 5.

Маленькая и большая амплитуда ядра фильтра одинаково плохо сказывается на его способности выделения границ, основанных на перепадах яркости. При слишком большой амплитуде фильтр окажется не чувствителен к мелким деталям, при слишком маленькой перестанет быть инвариантен к освещенности. Оптимальным видится использование параметров 2 и 3. Учитывая соображения по поводу схожести работы фильтра Габора и простых клеток зоны V1 зрительной системы, оптимальным будет параметр 2.

C_3 не оказывает особого влияния на ядро фильтра. Тем не менее, для получения именно периодического колебания с отрицательными и положительными зонами необходимо использовать 1 или 2 значения параметра, знак значения не имеет.

Для уменьшения зашумленности свертки, ядро фильтра должно иметь максимально близкую к нулю постоянную компоненту, определяемую как среднее значение ядра фильтра. В таблице 3 приведено сравнение различных комбинаций первых двух параметров (первая цифра обозначает параметр 1, вторая – 2)

Таблица 3: Сравнение параметров.

Вариант комбинации коэффициентов	Значение постоянной компоненты
2-2	2,726183204E-05
3-2	5,452366409E-05
1-2	8,564557128E-05
4-2	1,536646879E-04
5-2	4,330750086E-04
6-2	5,447271355E-04
2-1	9,161594046E-04
2-5	1,513560835E-03
3-1	1,832318809E-03
2-3	2,153531600E-03
1-1	2,878199655E-03
3-5	3,027121671E-03
3-3	4,307063200E-03
1-5	4,754991602E-03
1-3	6,765519053E-03
4-1	7,303063589E-03
2-4	8,071898710E-03
4-3	8,583319746E-03
3-4	1,614379742E-02
4-5	1,706274567E-02
5-3	1,710524653E-02
4-4	1,815117774E-02
5-4	2,040816327E-02
1-4	2,535861769E-02
6-1	2,588868636E-02
5-1	2,910778273E-02
6-3	3,042707627E-02
6-5	6,048585854E-02
6-4	6,434424978E-02
5-5	9,617627615E-02

Как можно заметить, наиболее близкие к нулю значения постоянной компоненты имеет ядро фильтра с коэффициентом $C_2 = \frac{1}{2}$. Параметр $C_1 = \frac{F^2}{\pi \sigma_x \sigma_y}$ был предложен Курки [7], и его применение дает наименьшее значение постоянной компоненты, но, как говорилось выше, при росте λ ядро фильтра имеет очень маленькие значения, что негативно сказывается на селективной способности. Тоже самое касается коэффициентов 1, 3, 4. Таким образом, оптимальной является формула (2), предложенная Петковым [6].

4. Подбор параметров ядра фильтра

Зависимость между пространственной частотой и масштабом фильтра становится очевидной при использовании фильтра Габора. К настоящему времени сложилось несколько подходов к определению этой зависимости:

1. Постоянные соотношения
2. Пропускная способность
3. Алгоритмы оптимизации

Таблица 4: Сравнение параметров.

1.	$\frac{\sigma}{\lambda} = 0.16$	7.	$\frac{\sigma}{\lambda} = 0.975$
2.	$\frac{\sigma}{\lambda} = 0.4$	8.	$\frac{\sigma}{\lambda} = \sqrt{2} \approx 1.414$
3.	$\frac{\sigma}{\lambda} \approx 0.425$	9.	$\frac{\sigma}{\lambda} = 3\sqrt{\frac{\ln 2}{2}} \approx 1.766$
4.	$\frac{\sigma}{\lambda} = 0.5$	10.	$\frac{\sigma}{\lambda} = 3.5$
5.	$\frac{\sigma}{\lambda} = 0.7$	11.	$\frac{\sigma}{\lambda} = 2\pi \approx 6.283$
6.	$\frac{\sigma}{\lambda} = 0.8$		

Подбор оптимальных параметров фильтра для формирования банка фильтров некоторые авторы рассматривают как задачу статистической оптимизации. Методы [8, 9, 10] сходным образом используют оптимальную модель для мульти-текстурной сегментации, базирующуюся на вычислении энергии Габора. В данной работе эти методы рассматриваться не будут, так как сильно зависят от изображения, поданного на вход.

Пропускная способность для определения параметров фильтра впервые предложенная [11], вычисляется через половину пиковой магнитуды ядра.

$$\frac{\sigma}{\lambda} = \frac{1}{\pi} \sqrt{\frac{\ln 2 \cdot 2^B + 1}{2 \cdot 2^B - 1}} \quad (4)$$

Зачастую, принимают $B = 1$. В этом случае, $\frac{\sigma}{\lambda}$. В качестве показателя сравнения эффективности подбора параметров в настоящей работе применяется отношение сигнал-шум SNR, предложенное [12].

$$SNR = -\left(\frac{\mu_L}{\sigma_L}\right)^2 \quad (5)$$

где: μ_L – среднее значение результата свертки, σ_L – стандартное отклонение результата свертки.

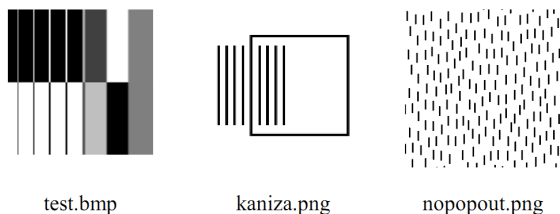


Рис. 1: Изображения для сравнительного анализа.

В таблице 3 приведено сравнение рассчитанного для 3 изображений значения SNR. Свертка произведена с ядром, рассчитанным по формуле 2 ($\gamma = 1, \lambda = 2, \theta = 0$) Изображения были специально

подобраны с большим количеством вертикальных элементов (рисунок 1).

Как видно из таблицы 5, наибольшие значения имеют соотношение $\frac{\sigma}{\lambda} = 0.5$ и соотношение, определяемое пропускной способностью.

Таблица 5: Сравнение значений для подбора $\frac{\sigma}{\lambda}$.

	SNR		
	test.bmp	kaniza.png	nopopout.png
0,16	-1,217	-2,825	-2,873
0,4	-0,617	-1,089	-1,024
0,424661	-0,472	-0,81	-0,759
0,5	-0,223	-0,373	-0,349
0,56	-0,237	-0,398	-0,373
0,7	-0,543	-0,965	-0,913
0,8	-0,739	-1,403	-1,34
0,975	-0,932	-1,947	-1,89
1,414214	-1,082	-2,521	-2,501
1,766115	-1,118	-2,686	-2,684
3,5	-1,156	-2,883	-2,908
6,283185	-1,164	-2,925	-2,957

5. Заключение

Приведенная структура слоя выделения краев может помочь в дальнейшем построении нейросетей компьютерного зрения, так как предоставляет информацию о границах сегментах и возможных признаках объекта. Приведенные типы рецептивных полей полностью соответствуют возможным конфигурациям границ. Косвенно, атомарность данных типов подтверждается возможностью составить из них все признаки Хаара.

Оптимальной для использования является формула, предложенная Петковым. Это подтверждает расчет постоянной компоненты и аналитическое сравнение влияния различных значений длины волны на форму графика ядра. Также было приведено описание способов подбора соотношения длины волны и масштаба фильтра. Применение показателя сигнал-шум позволило говорить об оптимальности $\frac{\sigma}{\lambda} = 0.5$ и использовании пропускной способности.

Литература

- [1] Riesenhuber, M., Poggio T.; Hierarchical models of object recognition in cortex. Nature neuroscience, 2 (11), 1019–1025 (1999)

- [2] Franklin S.W., Rajan S.E.: Retinal vessel segmentation employing ANN technique by Gabor andmoment invariants-based features. *Applied Soft Computing*, 22, 94-100 (2014)
- [3] Oh S., Choe Y.: Segmentation of textures defined on flat vs. layered surfaces using neural networks: Comparison of 2D vs. 3D representations. *Neurocomputing*, 70, 2245-2255 (2007)
- [4] De A., Guo Ch.: An adaptive vector quantization approach for image segmentation based on SOM network. *Neurocomputing*, 149, 48-58 (2015)
- [5] Gao Ch., Zhou D., Guo Y.: Automatic iterative algorithm for image segmentation using a modified pulse-coupled neural network. *Neurocomputing*, 119, 332-338 (2013)
- [6] Petkov, N.: Biologically motivated computationally intensive approaches to image pattern recognition, *Future Generation Computer Systems*, 11 (4-5), 451-465 (1995)
- [7] Kyrki, V.: Local and global feature extraction for invariant object recognition. Ph.D. thesis. Lappeenranta University of Technology (2002)
- [8] He, F., Liu, Y., Zhu, X., et al.: Multiple local feature representations and their fusion based on an SVR model for iris recognition using optimized Gabor filters. *EURASIP Journal on Advances in Signal Processing*, 2014:95 (June 2014)
- [9] Li, M., Staunton, R.C.: Optimum Gabor filter design and local binary patterns for texture segmentation. *Pattern Recognition Letters*, 29(5), 664-672 (2008)
- [10] Tsai, D.-M., Wu, S.-K., Chen, M.-Ch.: Optimal Gabor filter design for texture segmentation using stochastic optimization. *Image and Vision Computing*, 19(5), 299-316 (2001)
- [11] Bovik, A.C., Clark, M., Geisler, W.S.: Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Machine Intell.*, 12(1), 55-73 (1990)
- [12] Bodnarova, A., Bennamoun, M., Latham, S.: Optimal Gabor filters for textile flaw detection. *Pattern Recognition*, 35(12), 2973-2991 (2002)