



Politechnika Wrocławská

Wydział Informatyki i Zarządzania

kierunek studiów: Informatyka

specjalność: Inżynieria oprogramowania

Praca dyplomowa - magisterska

**ROZPOZNAWANIE OBRAZÓW PODOBNYCH PRZY
UŻYCIU PODEJŚCIA WZOROWANEGO NA KORZE
WZROKOWEJ**

Adam Puchalski

słowa kluczowe:

kora wzrokowa

rozpoznawanie obrazów

sztuczna inteligencja

krótkie streszczenie:

Praca obejmuje opis kilku podejść – w tym jedno autora – w rozpoznawaniu obrazów, które bazują na korze wzrokowej ssaków.

Promotor:
	<i>imię i nazwisko</i>	<i>ocena</i>	<i>podpis</i>

Wrocław 2010

Niniejszy dokument został złożony w systemie L^AT_EX.

Spis treści

Wprowadzenie	1
Rozdział 1. Biologiczne podstawy pracy	3
1.1. Mózg człowieka	3
1.2. Wydzielone obszary w mózgu	3
1.3. Kora wzrokowa	4
1.4. O wzroku słów kilka	5
1.5. Struktura kory wzrokowej	5
1.5.1. Pierwszorzędowa kora wzrokowa – V1	5
1.5.2. Pozostałe warstwy kory wzrokowej	7
Rozdział 2. Modele kory wzrokowej z zastosowaniem praktycznym	9
2.1. Ekstrakcja i integracja konturów	9
2.1.1. Ekstrakcja konturów z obrazu	9
2.1.2. Tablica cech wizualnych – VFA	12
2.1.3. Hiperkolumna	14
2.1.4. Integracja konturów w pierwszorzędowej korze wzrokowej	16
2.1.5. Integracja konturów modelem integrate-and-fire	19
2.1.6. Odzyskiwanie globalnych konturów z lokalnych cech	21
2.2. Pełny model obliczeniowy kory wzrokowej człowieka	24
2.2.1. Rozpoznawanie obiektów wywodzące się z podziału na komponenty	26
2.3. Podsumowanie	26
Dodatek A. Filtr Gabora	33
A.1. Teoria	33
A.1.1. Funkcja Gabora	34
A.1.2. Filtr Gabora	36
A.1.3. Przykład filtrowania	38
A.2. Dyskusja i wnioski	39
Spis rysunków	43
Spis tablic	45
Spis algorytmów	47
Bibliografia	49

Streszczenie

Abstrakt

Abstract

Abstract

Wprowadzenie

Rozpoznawanie obrazów jest jednym z najważniejszych zastosowań sztucznej inteligencji. Właściwie, aby można było mówić o sztucznej inteligencji konieczne jest uwzględnienie zaawansowanych technik rozpoznawania obrazów. Mimo, że zdecydowana większość problemów sztucznej inteligencji jest problemami trudnymi¹, to wydaje się, że na równi z rozpoznawaniem języka, jako podstawowe należy uznawać rozpoznawanie obrazów.

W tej pracy przedstawione będą podstawowe techniki stosowane w rozpoznawaniu obrazów, które w idei mają być komputerowym odpowiednikiem kory wzrokowej. Techniki te są mniej lub bardziej związane z dziedzinami rozpoznawania obrazów w grafice komputerowej i naukami o mózgu, które wcześniej wspomnianej kory wzrokowej dotyczą. Można by pokusić się o narysowanie odcinka na płaszczyźnie, którego końcami byłby te dwie dziedziny. Wtedy wszystkie istniejące projekty dotyczące rozpoznawania obrazów można by umieścić na tej płaszczyźnie. Były by one umiejscowione na tym odcinku w pewnej odległości, która określałaby powiązanie z konkretną dziedziną. Celem tej pracy jest znalezienie tego najlepszego punktu na tym odcinku, który będzie wykorzystywać zaawansowane techniki z dziedziny rozpoznawania obrazów i te zaobserwowane w korze wzrokowej. Zagadnienie to zostało sprowadzone do jednego wymiaru celowo, żeby pokazać skąd należy czerpać inspiracje, faktem jest jednak, że ciężko jest wyrazić złożoność tematu w jednym wymiarze. Aby lepiej przedstawić to co się dzieje w dziedzinie należałyby zdaniem autora skupić się na korze wzrokowej i dzieląc ją na elementy pokazać co, gdzie i w jaki sposób jest przedstawiane.

1. W literaturze angielskojęzycznej określa się to pojęciem AI-complete.

Rozdział 1

Biologiczne podstawy pracy

1.1. Mózg człowieka

Przy opisie mózgu człowieka wypada zacytować pewne porównanie, na jakie pokusił się David H. Hubel w swoim artykule z kwietnia 1978 roku [11]. Przedstawia on powody, przez które jest trudno mózg badać. Przede wszystkim skomplikowanie wiążące się z liczbą komórek z których się składa – 10^{10} , chociaż w tamtych czasach mówiło się nawet o 10^{11} . Dzisiaj mówi się o liczbie $2 \cdot 10^{10}$ [19] lub (w tej samej encyklopedii) $11 - 12 \cdot 10^9$. Wracając jednak do Hubela, liczba komórek w mózgu to nie jedyna bariera. Przywołuje on jako przykład wątrobę, która składa się jedynie z dwóch różnych komórek, które pracując wspólnie wykonując około 15 różnych funkcji, zakładając, że ich liczba może być większa. W mózgu natomiast szacuje się, że istnieje od 100 do 500 różnych komórek – w zależności od dokładności z jaką zostaną podzielone. Hubel nie zaryzykował szacowania liczby różnych funkcji mózgu. Wspomniał jednak, że dodatkowe skomplikowanie jest powodowane przez połączenia komórek ze sobą, które – jak można przypuszczać – przypominają gęstą sieć. W cytowanej wyżej encyklopedii, która jest podstawą kolejnych części tej pracy, szacuje się, że każda komórka jest powiązana z około 10^4 innymi co powoduje, że całkowita liczba połączeń synaptycznych może przekroczyć całkowitą liczbę cząsteczek występujących na ziemi. Przez co cały czas ludzki mózg jest dla naukowców zagadką i zarówno w 1978 roku jak i teraz więcej jest pytań jego dotyczących niż odpowiedzi.

1.2. Wydzielone obszary w mózgu

Mózg człowieka, mimo, że nie od samego początku uważany za organ odpowiadający za myślenie i przetwarzanie wszelkich informacji z otaczającego środowiska, był dla naukowców zawsze zagadką. To co dzisiaj wiadomo o mózgu jest efektem prac tysięcy naukowców, którzy podzieliли mózg na dziesiątki obszarów. Pierwszy taki podział był zaproponowany w 1907 roku przez Korbiniana Brodmanna. Podzielił on mózg na 47 obszarów [5] ze względu na różnice w strukturze warstwowej i komórkowej. Kora wzrokowa – podzielona na pierwotną (pierwszorzędową), drugorzędową oraz trzeciorzęgową – miała numery odpowiednio 17, 18 i 19. Dodatkowo utrwało się w literaturze

oznaczenie V1, V2 i V3, które odnosi się tylko do tych obszarów i również odpowiada polom 17, 18 i 19.

1.3. Kora wzrokowa

Zanim zostanie przybliżona struktura kory wzrokowej należy wspomnieć o tym w jaki sposób wiedza, którą posiadamy na temat tej części mózgu jest poszerzana. Niektóre badania prowadzone są z udziałem ludzi posiadających najogólniej uszkodzenia kory wzrokowej [7]. Do tych patologii należą:

- ślepotą – ta, która nie jest wynikiem wad wrodzonych oka, bądź innych, tylko jest spowodowana uszkodzeniami natury neurologicznej (czytając tekst napisany językiem Braille'a aktywowane są obszary kory wzrokowej)
- achromatopsią – agnozja barw) – niezdolność do rozpoznawania barw nawet jeśli oko jest zdolne do ich rozpoznawania
- akinetopsja – powszechnie nazywana ślepotą ruchu
- prosopagnozją – ślepotą twarzy [6]

Mając takiego osobnika można doszukiwać się szczegółów dotyczących jego niepełnosprawności, a co za tym idzie przyczyn, skutków, miejsca w którym się znajduje i innych. Wydobywając informacje o wielu patologiach z wielu przypadków dosyć prosto jest je uogólnić i dojść do wniosku, że nie tylko ludzie, ale i wszystkie ssaki mogą mieć podobną strukturę i działanie. Mając takie informacje badania można prowadzić na zwierzętach (myszach, kotach, małpach) co dalej rozwija tą naukę. Należy się jednak szczególna uwaga w kwestii interpretacji wyników badań mózgów małp, na co zwraca uwagę autor pracy [14].

Nie tylko z patologii kory wzrokowej można uzyskiwać informacje o tej części mózgu. Jeden z przykładów prezentowanych w encyklopedii [19] pokazuje kota z założoną do kory wzrokowej elektrodą. Jest ona przyłączona do oscyloskopu, aby w ten sposób wizualizowane wyniki przedstawiały impulsy, które są w danej części mózgu kory przesyłane. Dzięki temu zabiegowi możliwe jest badanie aktywności kory wzrokowej względem obrazów, jakie są kotu przedstawiane.

Wymieniona wcześniej prosopagnozja jest chorobą neurologiczną, której skutkiem – u osoby chorującej – jest niemożność rozpoznawania twarzy. Wykrycie takiego zaburzenia u człowieka spowodowało powstanie teorii, że istnieje specjalna część kory wzrokowej, która odpowiada za rozpoznanie twarzy w mózgu [6]. Posługując się techniką funkcjonalnego magnetycznego rezonansu jądrowego można stwierdzić nawet wiele więcej. Autorzy pracy [4] twierdzą, że w mózgu są specjalne struktury, odpowiedzialne za rozpoznanie ludzkiego ciała. Idąc tym tropem, autor pracy [8] badał obszary w korze wzrokowej, które uaktywniają się podczas pokazania badanemu pacjentowi różnych zdjęć. Na tej podstawie był w stanie określić miejsca aktywujące się na widok miejsc, obiektów, bądź twarzy. Okazało się też, że odnalezione w ten sposób obszary wykazują większą aktywność, jeśli zdjęcie przedstawia coś znanego niż coś, czego badany nie znał.

Cytowany wcześniej Grill-Spector badając ludzką korę wzrokową i jej reakcje na przedstawiane obrazy doszedł do kolejnych ciekawych wniosków. Okazuje się, że przy po-

kazywaniu obrazów człowiekowi, tak samo szybko jest on w stanie stwierdzić, czy na obrazie jest jakiś obiekt (detekcja) jak i to, do jakiej klasy on należy (kategoryzacja) [9]. Identyfikacja obiektu zawsze wymagała więcej czasu.

W ramach ciekawostki można też wspomnieć o obszarach w mózgu, które są wrażliwe na ruch obrotowy [15]. O co chodzi w tym określeniu najlepiej tłumaczy sposób, w jaki autorzy tej pracy do takiego wniosku doszli. Badali oni aktywność pewnych obszarów mózgu podczas jazdy po rondzie raz zgodnie, a innym razem przeciwnie do ruchu wskazówek zegara.

1.4. O wzroku słów kilka

Nie sposób opisać biologicznych podstaw rozpoznawania obrazów u człowieka nie zaczynając od oka, a przecież ono jest najbardziej namacalnym i pierwszym za razem elementem toru, którym informacja wizualna jest przekazywana do mózgu.

Oko wprawdzie nie spełnia specjalnie skomplikowanego zadania. Twierdzi się nawet, że oko jest elementem czysto mechanicznym, który odpowiednio przygotowuje wiązkę światła do momentu, w którym pada ono na siatkówkę. Na niej wyświetlany jest obraz widziany przez oko, jest on znacznie pomniejszony i obrócony. Siatkówka jest pewnego rodzaju błoną, wyposażoną w fotoreceptory. W nich następuje zmiana sygnału wizualnego na sygnał chemiczny. Informacja wizualna w postaci chemicznej jest przekazywana nerwem wzrokowym do ciała kolankowatego bocznego, a stamtąd przenoszona do pierwotnej kory wzrokowej.

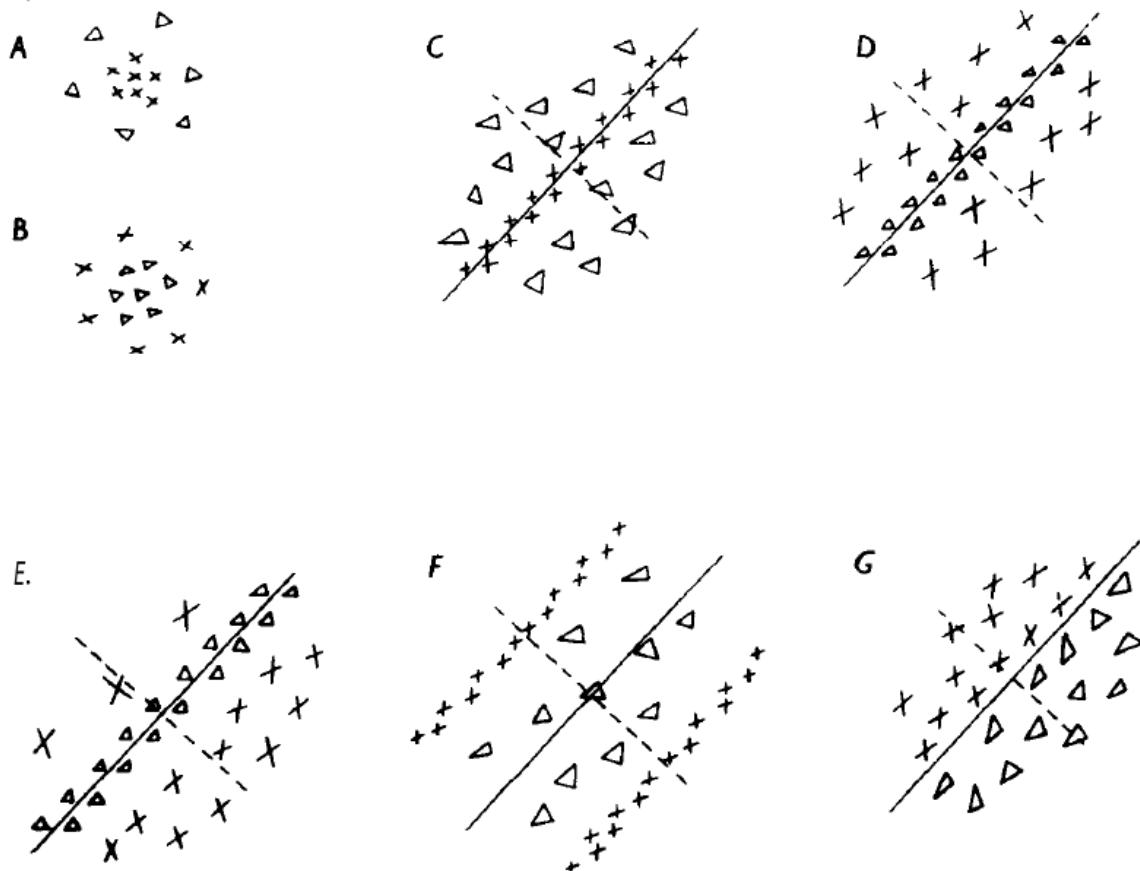
1.5. Struktura kory wzrokowej

Kora wzrokowa jest podzielona na obszary, które pełnią różne funkcje. Przedstawiają one jakby kolejne etapy przetwarzania bodźców wizualnych, jednak istnieją odstępstwa od tej reguły w postaci połączeń między-warstwowych. Mimo, że cały czas prowadzone są debaty co do zarówno liczby jak i granic wydzielanych obszarów w korze wzrokowej [14], naukowcy są zgodni co do wielości i zróżnicowania obszarów występujących w płacie potylicznym, w którym znajduje się kora wzrokowa. Prowadzone badania w tym zakresie są również zgodne co do początku neurologicznej części szlaku wzrokowego, którym jest pierwszorzędowa kora wzrokowa.

1.5.1. Pierwszorzędowa kora wzrokowa – V1

Inaczej nazywana pierwotną korą wzrokową, lub obszarem V1, odpowiada jej również numer 17 Brodmanna. Najważniejszych informacji dotyczących tej części kory wzrokowej dały prace Davida H. Hubela oraz Torstena N. Wiesela, które zostały w roku 1981 nagrodzone nagrodami Nobla w dziedzinie medycyny [12, 30].

Podstawowym faktem, jaki zaobserwowali naukowcy był fakt istnienia neuronów, których aktywność ściśle zależy od określonego obszaru na siatkówce (receptive



Rys. 1.1: Punkty A i B przedstawiaj± pola wzbudzeñ, neuronów z cia³a kolankowatego bocznego, co równie¿ odpowiada komórkom siatkówki oka. Punkty C-G przedstawiaj± pola wzbudzeñ neuronów pierwotnej kory wzrokowej, które s± wra¿liwe na istniejące w obrazie kreski, b dzie krawedzie pod okre¶lonym katem, lecz ró¿nych rozmiarów. Mo¿na sobie wyobrazi ,  e w korze wzrokowej dla bardzo wielu nachyleñ kąta s± podobne grupy neuronów [12].

field – RF). W ten sposób mo¿na by okre¶li  pole na siatkówce, którego wyjście stanowi wzbudzenie dla tego neuronu. Podobnie wiêc jak na siatkówce ka dy fotoreceptor odpowiada za fragment obrazu, tak w pierwotnej korze wzrokowej ka demu fragmentowi obrazu odpowiada pewna grupa neuronów. Dodatkowo badaj c wzbudzenie tych neuronów odkryli,  e konkretne obszary pierwotnej kory wzrokowej s y czu e na wyst puj ce w obrazie (odwzorowane na siatkówce) kreski (jasne i ciemne – rysunek 1.1 C-F) lub wyra ne krawedzie (rysunek 1.1 G) b  d ce w jego polu. Wzbudzenie neuronu dodatkowo  ci le zależy od kąta pod którym dane kreski si y znajdują oraz miejsca, w którym w polu wpływającym na wzbudzenie komórki si y znajduje. Naukowcy zdecydowali si  nazwa  te neurony komórkami złożonymi, najprawdopodobniej z powodu,  e na jeden neuron sk ada si o bardzo wiele komórek siatkówki, a przez to równie  cia a kolankowatego bocznego (których pola wzbudzeñ przedstawione s y na rysunku 1.1 A i B), przez co reprezentuje on pewien kompleks komórek. Co mo e nie do ko ca wydawa  si o oczywiste, ka demu fragmentowi obrazu odpowiada szereg neuronów, których bu-

dowa jest taka sama, a różnią się jedynie kątem kreski, na który są szczególnie wrażliwe.

Na takiej samej zasadzie określa się istnienie w pierwotnej korze wzrokowej neuronów, wrażliwych na wcześniej wspomniane kreski poruszające się w określonym kierunku jak i istnieniu konkretnego koloru we fragmencie obrazu. Dodatkowo okazało się, że pole wrażliwości neuronów – komórek złożonych – ma różne wielkości i pola te się wzajemnie pokrywają.

Naukowcy wprowadzili również pojęcie komórek hiper-złożonych, które stanowiły połączenie kilku komórek złożonych. Takie komórki udało się zlokalizować u zwierząt pokazując im kreski określonej długości.

Oprócz odkrycia elementów pierwotnej kory wzrokowej, naukowcy umiejscowili je w bardziej złożonych strukturach, które nazywane były kolumnami orientacyjnymi i miały konkretne miejsce w pierwotnej korze wzrokowej w mózgu.

1.5.2. Pozostałe warstwy kory wzrokowej

Przez cały czas, począwszy od odkryć Davida H. Hubela oraz Torstena N. Wiesela, prowadzone są badania, mające na celu uszczegółowienie całej kory wzrokowej jak i samego mózgu. Przeglądając osiągnięcia naukowców można stwierdzić, że istnieje ponad 30 różnych obszarów w mózgu, których sposób połączenia nie jest jednak do końca jasny. Przedstawianych jest wiele modeli kory wzrokowej, z których wiele jest szczegółowych, mających na celu reprezentację architektury mózgu – tego co i jak jest połączone wewnętrz [28]. Z perspektywy informatyka można by pokusić się o porównanie szczegółowych modeli mózgu do grafu, którego węzłami są obszary mózgu, a krawędziami połączenia między tymi obszarami. Ciągła praca naukowców na tym polu powoduje, że odnajdywane są nowe sposoby współdziałania obszarów mózgu, przez co graf ten nieuchronnie zmierza do grafu pełnego, potwierdzając tym samym skomplikowanie mózgu. Inne modele natomiast przedstawiają struktury zainspirowane nauką o korze wzrokowej, których celem jest mieć przede wszystkim zastosowanie praktyczne.

Nim autor skupi się na zastosowaniach praktycznych – które są ściśle związane z tematyką pracy – przedstawiony zostanie główny podział, na dwa strumienie informacji wzrokowej, których początkiem jest pierwszorzędowa kora wzrokowa [28].

Szlak brzuszny – ventral stream – odpowiedzialny za identyfikację widzianych obiektów

Szlak grzbietowy – dorsal stream – odpowiedzialny za relacje między obiektami oraz poruszanie się obiektów w przestrzeni

Mimo, że w związkach z dziedziną artykułach nie jest to powiedziane w prost, większość realizacji kory wzrokowej (we wspomnianych zastosowaniach praktycznych) opisuje szlak brzuszny, a więc samą identyfikację widzianych obiektów. Realizacja obu zadań nie jest prosta. Jeszcze trudniejsze jest powiązanie ich ze sobą. Marzeniem wszystkich projektantów kolejnych podejść w tej dziedzinie jest jednak stworzenie takiego systemu, który podobnie jak wzrok człowieka jest w stanie rozpoznawać, katego-

ryzować, . . . widziane obiekty.

Opis pozostałych – wszystkich z wyłączeniem pierwotnej – warstw kory wzrokowej można właściwie w tym momencie przerwać. Co było bowiem sygnalizowane wcześniej, naukowcy mają problem z określeniem struktury, pełnionych funkcji, umiejscowienia jak i połączeń ze sobą kolejnych – wyższych warstw kory wzrokowej.

Przedstawione poniżej architektury czerpią w mniejszym, bądź większym stopniu z przedstawianych w literaturze modeli funkcjonowania mózgu. Wspólną ich cechą jest jednak bezsprzecznie pierwotna kora wzrokowa.

Rozdział 2

Modele kory wzrokowej z zastosowaniem praktycznym

Z konstrukcją systemów rozpoznających obrazy, szczególnie tych, bazujących na odkryciach w korze wzrokowej są wciąż trudności. Wiadomo bowiem co się mniej więcej dzieje w głowie, jak informacja wizualna jest przekazywana i każdy może się domyślić jak ma wyglądać efekt rozpoznawania obrazów. W kwestii realizacji – zgodnie z wstępem do tej pracy – jest różnie. W każdej publikacji, która opisuje jedno z propozowanych rozwiązań przedstawiony jest punkt widzenia, rzadko opisujący zaawansowany etap przetwarzania obrazu. Najczęściej są to etapy początkowe, przedstawiające transformacje mające miejsce w pierwszorzędowej korze wzrokowej, a więc ekstrakcję i integrację konturów. Nie spotykane są konkretne propozycje rozwiązań problemów występujących w wyższych warstwach kory wzrokowej samodzielnie. Ciężko bowiem jest określić podstawy, bądź jakąś bazę, na której można by budować dalej. Co autorzy rozwiązań podkreślają bardzo często, problematyka jest już na poziomie odzyskiwania krawędzi z obrazów.

Wprawdzie „kompletne” modele wykorzystujące elementy kory wzrokowej w rozpoznawaniu obrazów też się pojawiają, nie sposób zacząć jednak inaczej jak właśnie od ekstrakcji i integracji samych konturów.

2.1. Ekstrakcja i integracja konturów

2.1.1. Ekstrakcja konturów z obrazu

Algorytmów na ekstrakcję konturów z obrazu jest bardzo wiele, na dziesiątki różnych sposobów można szukać krawędzi obiektów na nich się znajdujących. Mimo, że jest ich wiele, i wyniki ich działania są zadowalające – chociaż z entuzjazmem najczęściej wypowiadają się sami autorzy – to problem jest trochę gdzie indziej. Krawędzie odzyskane z obrazów klasycznymi algorytmami, wciąż są pikslami w obrazie, które wprawdzie dla ludzkiego oka stanowią lepsze lub gorsze odwzorowanie konturów, ale do dalszego przetwarzania muszą być dodatkowo przygotowane. Jest to spowodowane przede wszystkim wymaganiami podyktowanymi przez temat tej pracy, który zakłada

czerpanie z kory wzrokowej jako źródła inspiracji dla algorytmów. W związku z czym wymagane jest przetwarzanie otrzymanych obrazów przedstawiających krawędzie do struktur danych reprezentujących te krawędzie.

Przedstawiony więc będzie jeden z wielu algorytmów, który z obrazów odzyskuje kontury z obrazu cyfrowego.

DICONEX

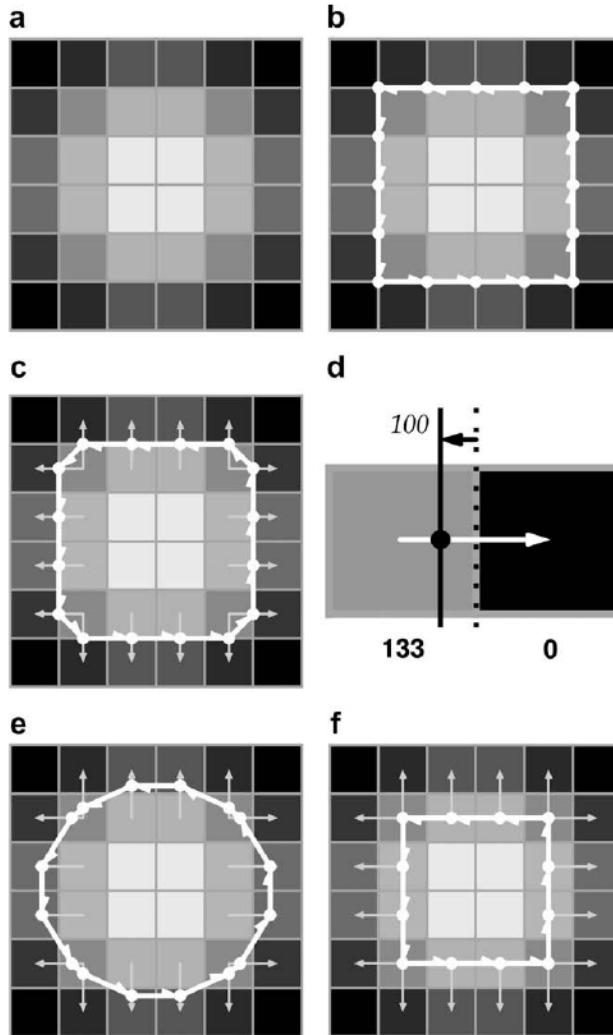
Algorytm – co podkreślają twórcy – jest perfekcyjny w sensie wyznaczania krawędzi obiektów jeśli przyjmie się następujące kryteria:

- obszary się wzajemnie nie przecinają,
- obszary są niepuste,
- nie ma ograniczeń na 2 wymiary,
- szybkość działania

Algorytm pracuje na obrazach monochromatycznych, więc wymaga się przekształcenia np. zdjęć do skali szarości.

Algorytm

Na początku algorytmu wymagane jest wybranie jednej – granicznej wartości jasności dla obrazu. W przykładzie z pracy [23], dla obrazu, którego jasność piksli jest z zakresu 0 – 155 została ona obrana na poziomie 100. Rysunek 2.1A przedstawia wejściowy obraz. W pierwszym kroku algorytmu, w obrazie wektorami konturowymi (ang. contour vector) otaczane są piksele o wartości jasności większej niż 100 – czyli jasne piksele – czego wynik przedstawiany jest na rysunku 2.1b. Warto zauważyć, że wspomniane wektory konturowe zawsze zwrócone są w taką stronę, że po lewej stronie mają otaczany obiekt – czyli w tym wypadku piksele o jasności większej niż 100. Zgodnie z tym algorytmem zawsze znalezione wektory będą ze sobą połączone na zasadzie początku jednego wektora z końcem drugiego wektora i nie będzie pustych obszarów. Jako kolejny krok algorytmu, wszystkie stworzone wektory zamieniane są na inne wektory, o tej samej orientacji, ale ich punkty zaczepienia znajdują się w środku wyznaczonego wcześniej wektora krawędziowego (konturowego), a kończą się w środku kolejnego wektora krawędziowego. Po zastosowaniu się do tego przekształcenia obraz otrzymany powinien wraz z krawędzią wyglądać tak jak na rysunku 2.1c. Dodatkowo na tym rysunku przedstawione są pewne wektory, które w cytowanej publikacji nazywane są wektorami zakresu (ang. range vector). Wektory te są umieszczone w obrazie w taki sposób, że łączą środki sąsiednich piksli i przechodzą przez każdy środek wektora konturowego a ich zwrot jest skierowany na zewnątrz otaczanego obszaru. Rysunek 2.1d przedstawia 2 piksele z obrazu, między którymi znajdował się środek wektora konturowego. W kolejnym kroku algorytmu punkt ten będzie przesunięty na wektorze zakresu w taki sposób, że obrana na samym początku wartość jasności (równa 100) będzie znajdować się na tym wektorze, gdy między środkowymi punktami piksli wartość ich jasności będzie liniowo interpolowana. Wynik takiego działania przedstawiony jest na rysunku 2.1e. Rysunek 2.1f przedstawia wynik transformacji tego poszerzonego konturu do konturu oznaczającego granicę szukanego obszaru. Jednak jak widać na przykładzie to przekształcenie jakby cofało wszystkie uogólnienia, zaokrąglenia i właściwie same przekształcenia, które zostały znalezione w obrazie 2.1c i 2.1e. Zdaniem

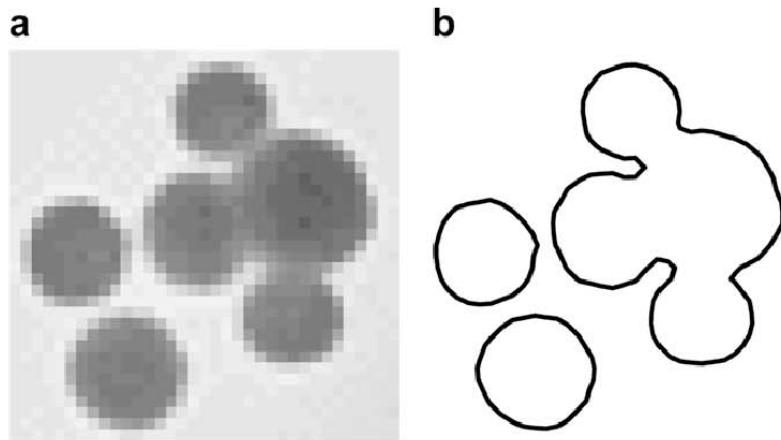


Rys. 2.1: Przykład działania algorytmu DICONEX na obrazie w skali szarości, o rozmiarach 6×6 [23].

autora najlepszy rezultat zaprezentowany jest na rysunku 2.1e.

O ile sam algorytm jest ciekawy, to został on tutaj zaprezentowany głównie z powodu na efekty, jakie można uzyskać w innych nietrywialnych obrazach, w szczególności tych, niewielkich rozmiarów. Rysunek 2.2 przedstawia wynik działania algorytmu na obrazie słabej rozdzielnosci. Wyraźnie widać, że kontury odtworzone przez algorytm są wyraźne i gładkie, co jest niewątpliwą zaletą.

Z zastosowaniem tego algorytmu jest jednak podstawowy problem, który polega na konieczności obrania wartości granicznej dla jasności grup pikseli w obrazie. Podobne problemy mają inne algorytmy odnajdujące krawędzie w obrazie. Można by się spierać, który jest lepszy, fakt jest jednak taki, że w oryginalnej postaci wymagają określania dodatkowych specyficznych parametrów.



Rys. 2.2: Wynik (b) działania algorytmu DICONEX na obrazie niewielkiej rozdzielczości (a) [23].

Trochę inaczej można podejść do tego zagadnienia metodami zaczerpniętymi wprost z grafiki komputerowej – filtrami. Dla obrazów binarnych (dwukolorowych) bardzo łatwo jest określić okno filtru, który miałby odzyskać same krawędzie z obszarów przedstawionych w obrazie [27]. Dla obrazów w skali szarości można posłużyć się bardziej zaawansowanymi filtrami, w tym na przykład filtrem Gabora, o którym dokładniej w części A.

2.1.2. Tablica cech wizualnych – VFA

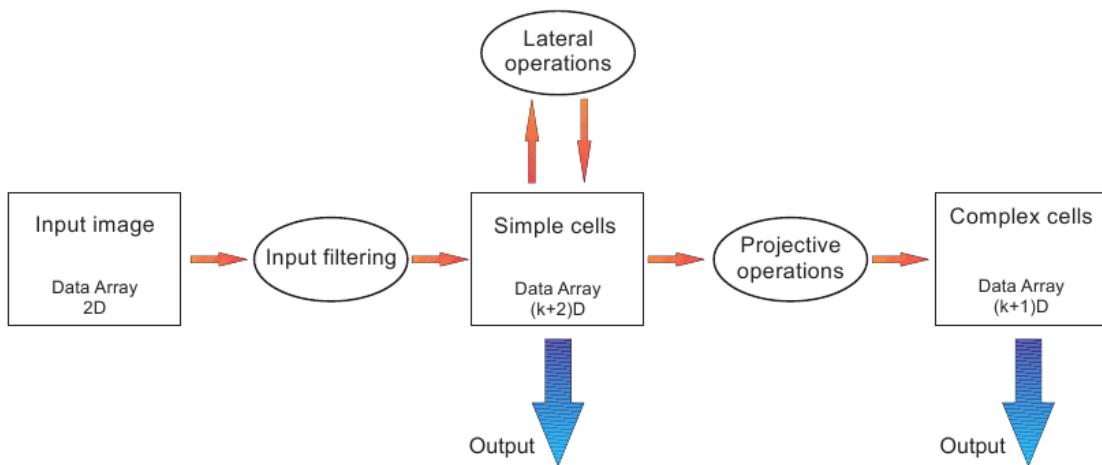
VFA – z angielskiego Visual Feature Array – można to tłumaczyć jako tablica cech wizualnych. Autorom prac [2, 20, 21] spodobała się idea zaproponowana przez wspomnianych wcześniej noblistów i zaproponowali model struktury powiązanych kolumn orientacyjnych, którą nazwali właśnie VFA.

Rysunek 2.3 przedstawia zaproponowany model tablicy cech wizualnych, w których w prostokąt oznaczony jako Input Image przedstawia obraz wejściowy. Jest on filtrowany i trafia do drugiego prostokąta, który reprezentuje „proste komórki” w korze wzrokowej. Na tym etapie następuje integracja konturów. Gdy kontury zostaną zintegrowane, trafiają do etapu składania wyodrębnionych krawędzi w większe struktury.

Filtrowanie

Przy tym etapie – wyodrębniania konturów – autorzy stosują dosyć dobrze rozpoznawany w grafice komputerowej filtr Gabora (załącznik A). Okazuje się, że natura w korze wzrokowej realizuje coś podobnego do tego matematycznego modelu, co jest opisane dosyć dobrze w pracy [13]. Świadczy to o tym, że zastosowanie filtra Gabora jest bardzo dobrym krokiem ku celu, którym jest wykorzystanie zalet kory wzrokowej w rozpoznawaniu obrazów.

Integracja konturów



Rys. 2.3: Model tablicy cech wizualnych, prostokąty symbolizują struktury danych, elipsy natomiast transformacje [20].

Na rysunku 2.3 opisana jako „Lateral operations”¹ integracja konturów, polega na tym, że neurony odpowiedzialne za krawędzie występujące w obrazie pod odpowiednim kątem mają fizyczne połączenie z neuronami odpowiedzialnymi za sąsiednie krawędzie, przez co wpływają na siebie nawzajem.

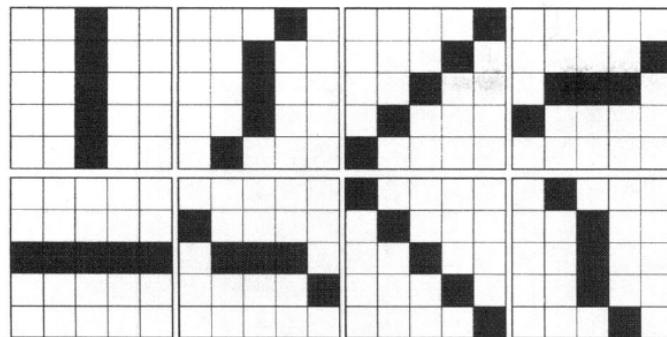
Autorzy założyli, że każda krawędź występująca w obrazie, a otrzymana w momencie ekstrakcji konturów miała przypisaną pewną wartość, która prezentuje jej istotność w obrazie. Wartość ta jest modyfikowana na skutek współdziałania neuronu odpowiedzialnego za tą krawędź z innymi, które z nim sąsiadują. Autorzy uznali, że dobrym podejściem będzie tutaj kombinacja liniowa istotności sąsiednich krawędzi w obrazie.

Łączenie krawędzi i przedstawienie rezultatów

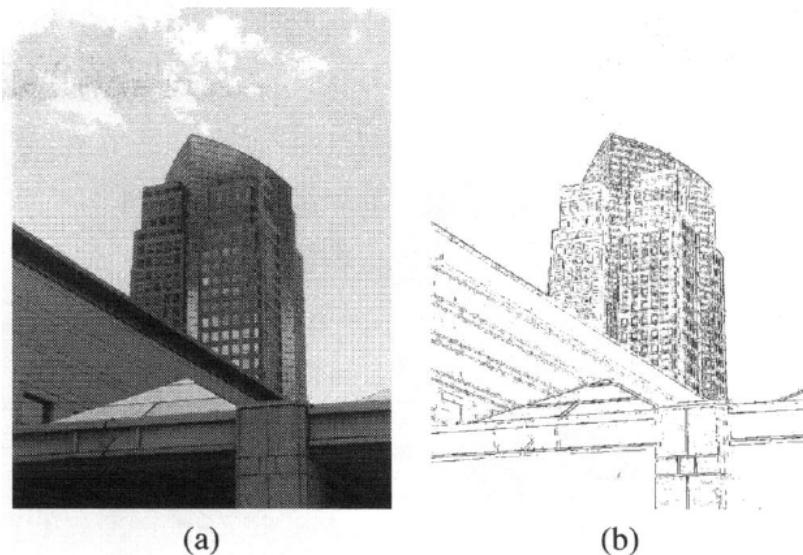
Mimo, że autorzy dosyć dobrze przedstawili samą teorię potrzebną do realizacji poprzednich etapów, to budowanie cech złożonych – które jest efektem łączenia krawędzi w większe struktury – jak i szczegóły dotyczące wcześniejszych etapów jest traktowana po macoszemu – o czym świadczy brak takich informacji w cytowanych pracach. Można odnieść wrażenie, że przedstawione są tutaj jedynie pomysły, a realizacja jest pozostawiona na uboczu.

Otrzymywane wyniki również specjalnie nie zachwycają, ponieważ nietrywialny przykład zastosowania jest tylko jeden, zaczerpnięty z pracy [21].

1. Lateral Geniculate Nucleus – z języka angielskiego – ciało kolankowate boczne. Jest jednym z elementów szlaku wzrokowego – którym informacja wizualna jest przenoszona do – między innymi – pierwotnej kory wzrokowej, chociaż połączenie jest dużo więcej niż tylko to z pierwszorzędową korą wzrokową. W sensie modelu określenie to ma symbolizować połączenia neuronów sąsiadujących będących w pewnym otoczeniu ze sobą.



Rys. 2.4: Krawędzie z modelu tablicy cech wizualnych, mające reprezentować krawędzie wyodrębnione w obrazie [21].

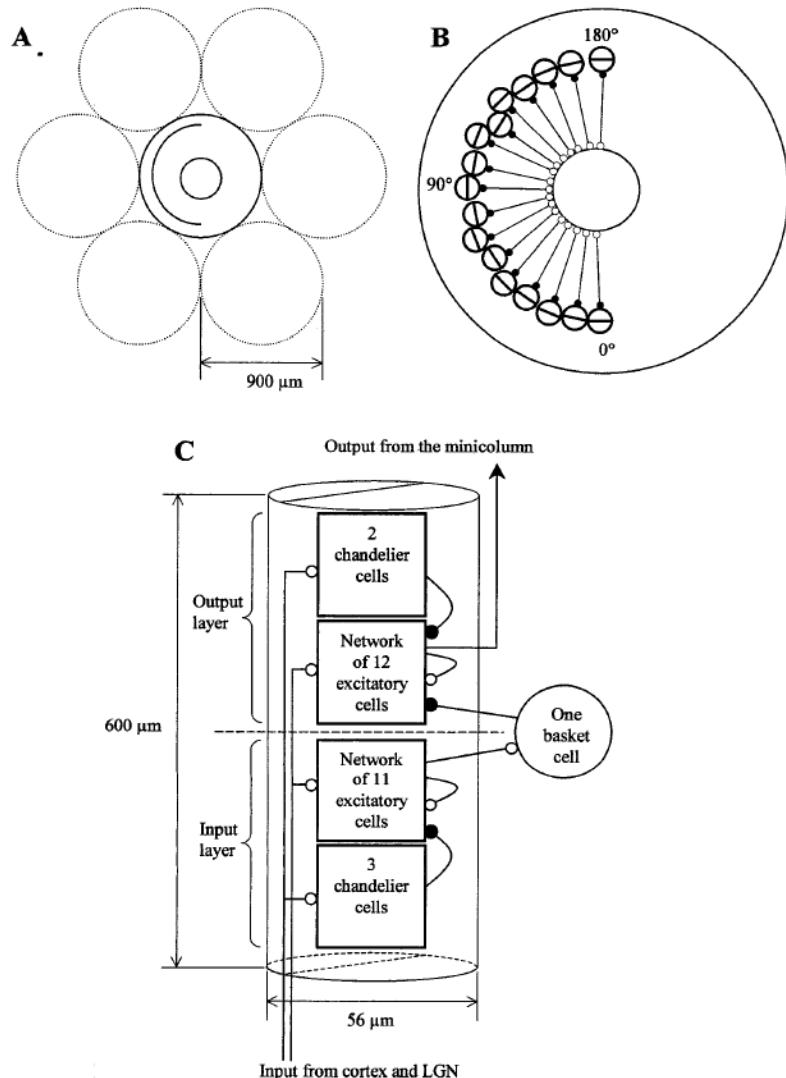


Rys. 2.5: Rezultaty pracy modelu z obrazem: a) obraz wejściowy, b) wyjście z modelu [21].

2.1.3. Hiperkolumna

Patrząc na odzyskiwanie konturów z obrazu różnymi metodami w końcu trafi się na perspektywę, która będzie wywodziła się bezpośrednio z kory wzrokowej. Praca [3] przedstawia szczególny punkt widzenia w kwestii szukania krawędzi w obrazie, można w niej przeczytać o hiperkolumnie, o której wspomniane było w części 1.5. Praca ta mimo, że opisuje tylko pewien model, może przedstawać dobry pomysł na realizację zadania odzyskiwania konturów w postaci zbliżonej do tych, które są w korze wzrokowej człowieka.

Autorzy proponują, aby każdemu obszarowi w obrazie odpowiadała pewna hiperkolumna (rysunek 2.6A) składająca się z 17 różnie ukierunkowanych krawędzi (rysunek



Rys. 2.6: Model hiperkolumny, zaproponowany przez autorów pracy [3], powstały na podstawie badań kory wzrokowej kota.

2.6B), które są połączone w ramach tej kolumny w sieć BCPNN² mającej za zadanie wybrać spośród 17 krawędzi będących pod różnymi kątami tą najbardziej odpowiednią. Kolumna ta składa się z dwóch warstw (rysunek 2.6C), wejściowej i wyjściowej, w których łącznie 28 neuronów. Nie ma fizycznych połączeń między tymi dwoma warstwami innych niż poprzez jedną komórkę którą autorzy nazwali koszykiem³. Komórka ta pełni rolę inhibitora – czyli powoduje osłabienie sygnału wyjściowego z neuronów⁴. Połączona jest natomiast z ekscytatorami, czyli komórkami, które powodują wzmacnianie sygnału wyjściowego z neuronów. Wszystkie neurony warstwy wejściowej są

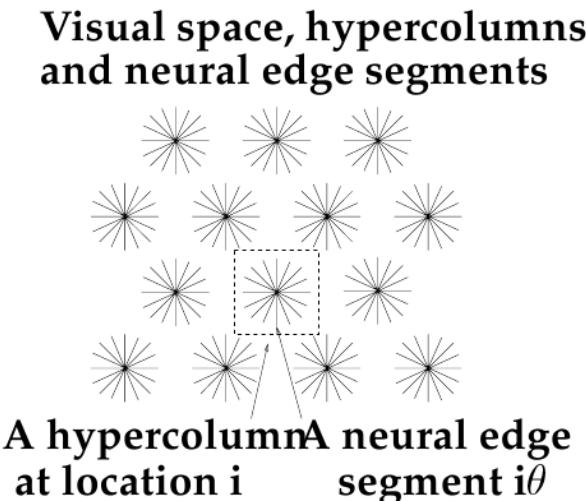
- 2. Bayesian Confidence Propagation Neural Network – Sieć neuronowa zbudowana na podstawie sieci Bayesowskiej.
- 3. Z języka angielskiego – basket cell.
- 4. W literaturze angielskojęzycznej spotyka się określenie firing rate – które nie do końca szczegółowo, ale jest w tej pracy tłumaczone na siłę sygnału wyjściowego. Jest to jednak tylko przeniesienie języka sztucznej inteligencji do stricte biologicznego opisu.

połączone z komórką koszykową, przez co mają na nią bezpośredni wpływ. Podobnie z komórkami warstwy wyjściowej, których wejścia sterowane są jedynie z łączącej obie warstwy komórki koszykowej. Dodatkowo w każdej warstwie kolumny znajdują się inhibitory, których zadaniem jest wpływanie na eksyktatory. Widać wyraźnie, że wyjściem z hiperkolumny jest sygnał z eksyktatorów warstwy wyjściowej. Natomiast samo wejście jest przekazywane z ciała kolankowatego bocznego.

Opis ten można by kontynuować wprowadzając więcej czysto biologicznych szczegółów. Autor tego dokumentu chciał jednak wprowadzając ten model pokazać z jednej strony, jak interpretowane mogą być odkrycia opisywane w części 1.5 tej pracy, z drugiej natomiast przedstawić pewną strukturę, która jest wykorzystywana w kilku kolejnych podejściach stosowanych głównie do symulacji działania pierwszorzędowej kory wzrokowej.

2.1.4. Integracja konturów w pierwszorzędowej korze wzrokowej

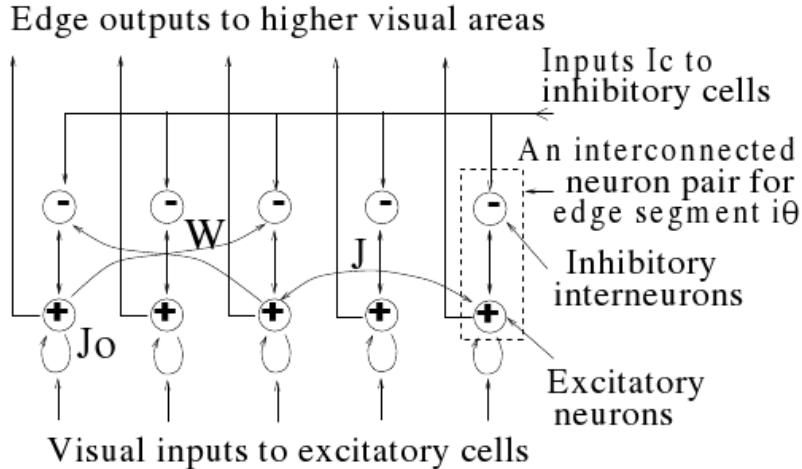
Aby zaczerpnąć z prezentowanych wcześniej technik 2.1.1 oraz 2.1.3 został stworzony inny model ekstrakcji i integracji konturów.



Rys. 2.7: Przestrzeń hiperkolumn w modelu integracji konturów [16].

Podstawowym elementem modelu integracji konturów jest hiperkolumna – która zgodnie z tym co zostało zaobserwowane w korze wzrokowej kota – jest ona umieszczona w heksagonalnej przestrzeni, co znaczy, że każda z nich 6 sąsiadujących. Składa się ona z segmentów krawędziowych⁵, a każdy segment może zostać określony numerem kolumny i w przypadku jeśli wszystkie są ponumerowane oraz kątem θ który odpowiada za nachylenie krawędzi w segmencie. W każdym segmencie znajduje się para neuronów, inhibitor i eksyktator, które są ze sobą połączone. Gdy w części obrazu objętej obszarem i znajduje się krawędź zorientowana pod kątem β oraz siłą/intensywnością

5. Tłumaczenie z języka angielskiego – edge segment.



Rys. 2.8: Architektura częściowa sieci prezentującej podstawowe elementy składowe modelu z pracy [16].

$\hat{I}_{i\beta}$ (rysunek 2.7), wtedy segment otrzymuje na wejście sygnał $I_{i\beta}$ który można wyliczyć z poniższego wzoru.

$$I_{i\beta} = \hat{I}_{i\beta} \phi(\theta - \beta)$$

gdzie

$$\phi(\theta - \beta) = e^{-|\theta - \beta|/(\phi/8)}$$

jest pewną funkcją odchylenia. Rzecz jasna segmenty poza hiperkolumną nie dostają żadnego sygnału.

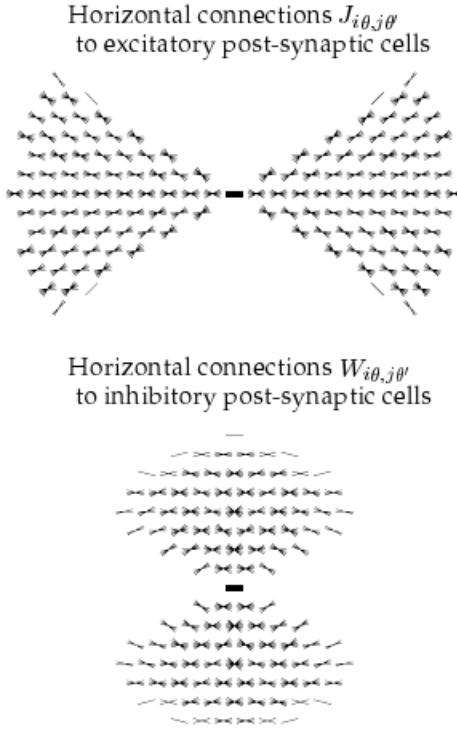
Określa się też parametr nazywany barierą membranową oznaczaną $x_{i\theta}$ dla ekscytatora oraz $y_{i\theta}$ dla inhibitora oraz funkcje wyjścia dla każdego z neuronów $g_x(x_{i\theta})$ oraz $g_y(y_{i\theta})$. Architektura tej sieci jest zaprezentowana na rysunku 2.8.

Wyniki działania modelu – wprawdzie skąpe, są przedstawione na rysunku 2.10.

Autor cytowanej pracy zwraca uwagę na połączenia między-segmentowe. Wyraźnie widać je na schemacie prezentującym architekturę sieci. Co jest istotne, są one realizowane jedynie na 2 sposoby: między ekscytatorami różnych segmentów oraz między ekscytatorem i inhibitorem z różnych segmentów. Nie ma połączeń między inhibitorami, nawet na siebie.

Sens sieci jest taki, aby ona się omiatała przez jakiś czas i ustalała swój stan. Konieczne jest bowiem propagowanie wyników – sygnałów wyjściowych po raz kolejny do wejścia neuronów oraz przeliczanie od nowa. Bardzo dobrze jest to przedstawione na rysunku 2.8.

Aby jednak móc lepiej zaprezentować działanie całości potrzeba trochę formalizmów,



Rys. 2.9: Pole mające wpływ na konkretny jeden segment, w zależności od funkcji odległości $J_{i\theta,j\theta'}$ oraz $W_{i\theta,j\theta'}$ [16].

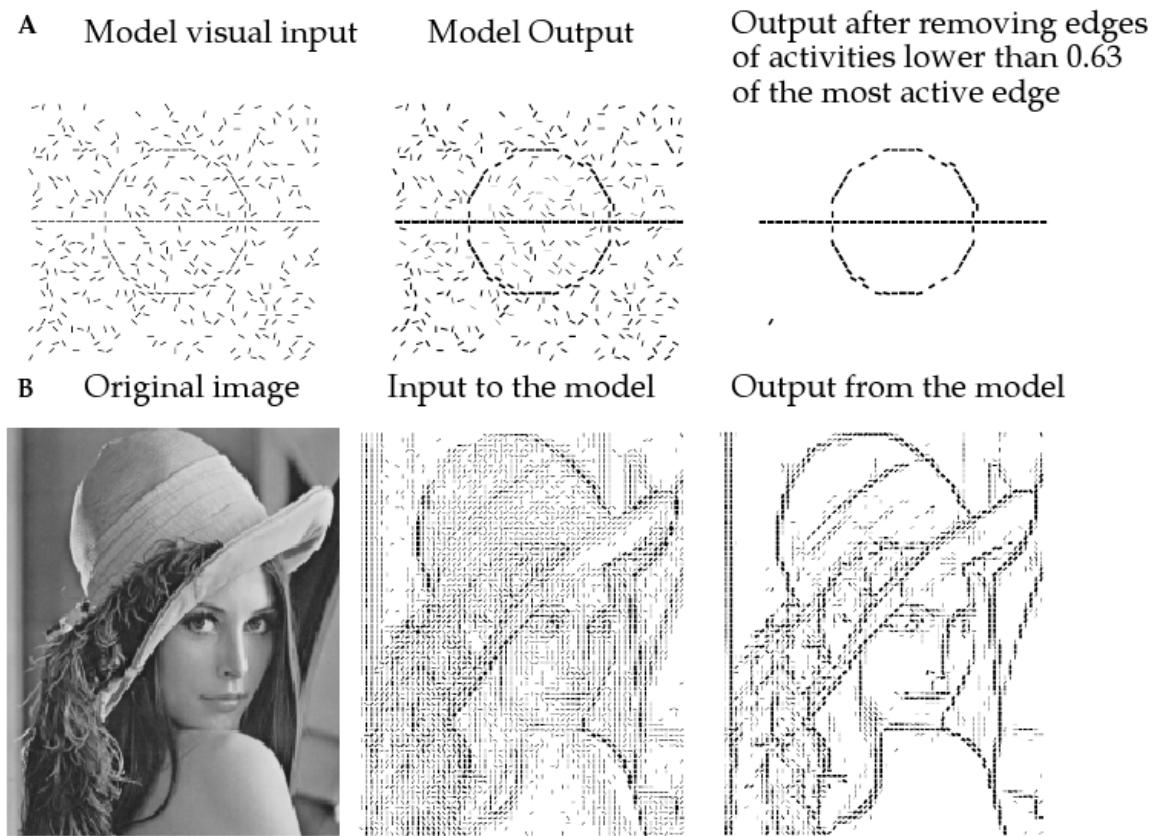
których w pracy [16] nie brakuje.

$$\begin{aligned} \dot{x}_{i\theta} = & -\alpha_x x_{i\theta} - \sum_{\Delta\theta} \psi(\Delta\theta) g_y(y_{i\theta+\Delta\theta}) + \\ & + J_0 g_x(x_{i\theta}) + \sum_{j \neq i, \theta'} J_{i\theta,j\theta'} g_x(x_{j\theta'}) + I_{i\theta} + I_0 \end{aligned} \quad (2.1)$$

$$\dot{y}_{i\theta} = -\alpha_y y_{i\theta} - g_x(x_{i\theta}) + \sum_{j \neq i, \theta'} W_{i\theta,j\theta'} g_x(x_{j\theta'}) + I_c \quad (2.2)$$

Wzory 2.1 oraz 2.2 prezentują wyjściowe stany w tak zwanym momencie ' $t + 1$ '. Wyступające we wzorach $J_{i\theta,j\theta'}$ oraz $W_{i\theta,j\theta'}$ są w pewnym sensie funkcjami odległości $j\theta'$ od $i\theta$. Funkcje te są pewnej nieprzystępnej postaci – wyznaczonej doświadczalnie i nie będą tutaj umieszczane, konieczne jest jednak pokazanie sensu tych funkcji. Można go zaprezentować określając pole, jakie ma znaczenie biorąc pod uwagę wpływ sąsiednich pól, na jedno – konkretne. Przedstawione ono jest na rysunku 2.9.

Omawiając to podejście należy wspomnieć o uwagach autorki publikacji [16] z której zostały zaczerpnięte przykłady. We wstępie swojej pracy zwraca Ona uwagę na problematykę etapu ekstrakcji konturów z obrazu. Wymienia kilka różnych algorytmów, na które mają związek z tym, jak kontury odzyskiwane są w korze wzrokowej. Zwraca



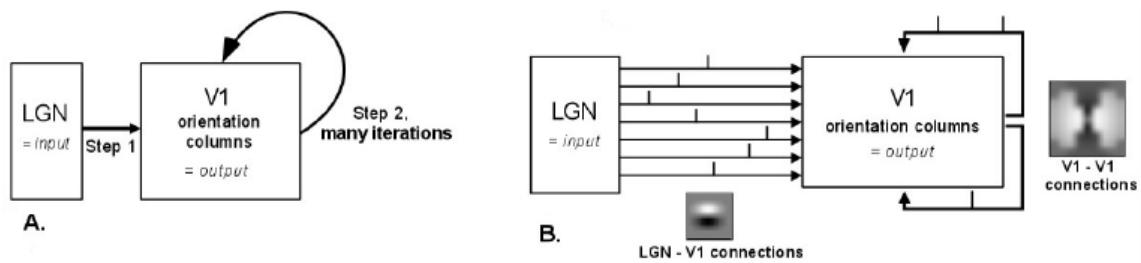
Rys. 2.10: Wyniki działania opisanego wcześniej algorytmu. Praca przedstawia zdjęcie najczęściej poddawane obróbce cyfrowej – Lenę. Wyraźnie widać, że krawędzie zostały odzyskane w praktycznie doskonały sposób [16].

uwagę jednak na to, że nawet bez ograniczeń wynikających z zastosowania podejścia wzorowanego na pierwszorzędowej korze wzrokowej jest to zagadnienie trudne. Wspomniany w części 2.1.1 algorytm na proste i skuteczne odzyskiwanie krawędzi z obrazu, wprawdzie nie był zastosowany w opisywanej pracy, ale inne algorytmy tego pokroju były przez autorkę testowane. Odnosi się bowiem Ona do nich jako do algorytmów mających problem. Najczęściej wynika on z występowania parametrów, które wymagają dostrojenia przez użytkownika, przez co do dobrej symulacji ekstrakcji konturów nie mogą być zastosowane.

Autorka stawia nawet pytanie, czy jest to osiągalne za pomocą modelowania obszaru V1 mózgu, czy może to wymagać wpływu wyższych warstw kory wzrokowej. Opisywane wcześniej biologiczne podstawy funkcjonowania kory wzrokowej wskazują, że może to być bardzo prawdopodobne.

2.1.5. Integracja konturów modelem integrate-and-fire

Autorzy pracy [29] zauważyl pewien mankament zwłaszcza wydajnościowy w modelu przedstawionym choćby w poprzedniej części – 2.1.4. Podstawową jego cechą były dziesiątki połączeń miedzy neuronami z hiperkolumn, przez co przeliczanie się całości



Rys. 2.11: Przedstawienie porównania modeli integracji konturów w pierwszorzędowej korze wzrokowej [29].

trwało jakiś czas. Na podobnej zasadzie autorzy przytaczają też inne publikacje.

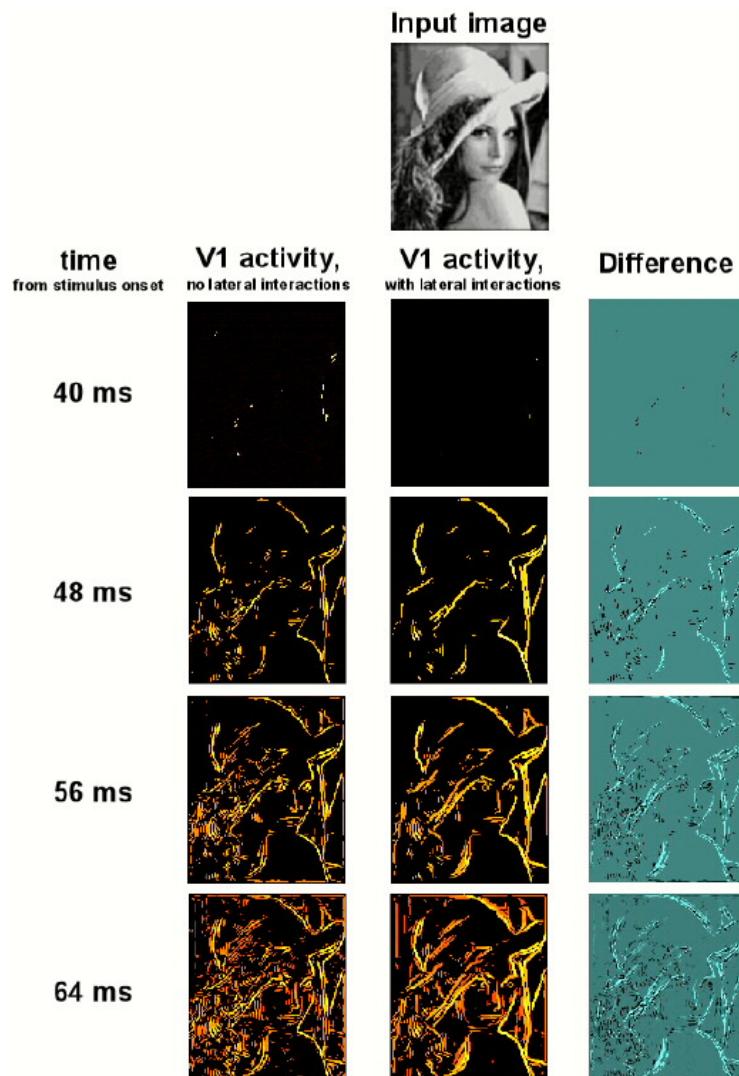
Głównym argumentem – który zdaniem autora tego dokumentu jest śmieszny – który stał za wnioskiem, że wyliczone modele są złe, był czas ich działania. Porównując je bowiem z błyskawicznym⁶ przetwarzaniem obrazów w korze wzrokowej człowieka, bądź małpy, muszą one być w sprzeczności z całą biologią, która za tym stoi. Nie podejmowano jednak argumentu, że w mózgu absolutnie wszystko jest „przeliczane” równolegle i nie ma w nim miejsca na iteracje znane z komputerowych implementacji sieci neuronowych. Nie mniej jednak proponuje się w tej pracy specyficzne podejście bazujące na modelu „sumuj i strzelaj”⁷. Zaproponowany model przedstawia rysunek 2.11B. Autorzy porównują go do proponowanych w innych pracach iteracyjnych podejścia, które przedstawione są na rysunku 2.11A. Na pierwszy rzut oka nie widać szczególnie różnych, podstawowa wynikająca z zastosowania podejścia Lapicque'a jest jednak – zdaniem autorów – kluczowa.

W modelu widać wyraźnie sprzężenie zwrotne i mimo, że są one krytykowane w cytowanych wcześniej pracach, to autorzy podkreślają, że to nie jest problem samych sprzężeń. Chodzi, raczej o to, że w klasycznych modelach sprzężenie zwrotne działa po przeliczeniu się wszystkich neuronów, tutaj jednak, przez wzgląd na obserwacje kory wzrokowej kota sprzężenie zwrotne cały czas ma wpływ na neurony warstwy V1. Całą implementację upraszcza dodatkowo fakt, że przez połączenia prowadzone są jedynie pewne piki, a same neurony zliczają je i na tej podstawie określają swój stan. Nie potrzeba tutaj przesyłania ciągłych sygnałów.

Rysunek 2.12 przedstawia otrzymywane w ten sposób rezultaty, które zgodnie z poprzednim akapitem zostały sobie przeciwstawione. Po lewej stronie rezultaty prezentują wyjście z modelu, w którym połączenie między elementami pierwszorzędowej kory wzrokowej nie ma, środkowe obrazy przedstawiają te same wyniki modelu, z różnicą na występowanie połączeń między elementami V1. Prawa kolumna pokazuje różnice między opisanymi rezultatami. Tak jak i w dziesiątkach książek poświęconych grafice komputerowej i tutaj obrazem do testowania była Lena. Podstawowym parametrem

6. W publikacji jest to określone angielskim słowem rapid.

7. Tłumaczenie z języka angielskiego frazy integrate-and-fire, którą można nazwać inaczej modelem Lapicque'a[1].



Rys. 2.12: Porównanie wyników otrzymywanych z modelu „integrate-and-fire” ze sprzężeniem i bez [29].

modelu jest czas, jaki się poświęca przeliczeniom samego modelu. Otrzymywane rezultaty przedstawiają przede wszystkim zalety korzystania z połączeń między elementami obszaru V1 jak i samego pomysłu na realizację integracji konturów w korze wzrokowej.

2.1.6. Odzyskiwanie globalnych konturów z lokalnych cech

W celu przeprowadzenia integracji konturów wcale nie trzeba budować sieci neuronowej tak jak to było pokazane w poprzednich częściach tej pracy. Autorzy pracy [10] przedstawiają trochę inny punkt widzenia. Jako pierwsze wprowadzają pojęcie pola rozszerzenia⁸ jako wzorca reprezentującego rodzinę gładkich krzywych.

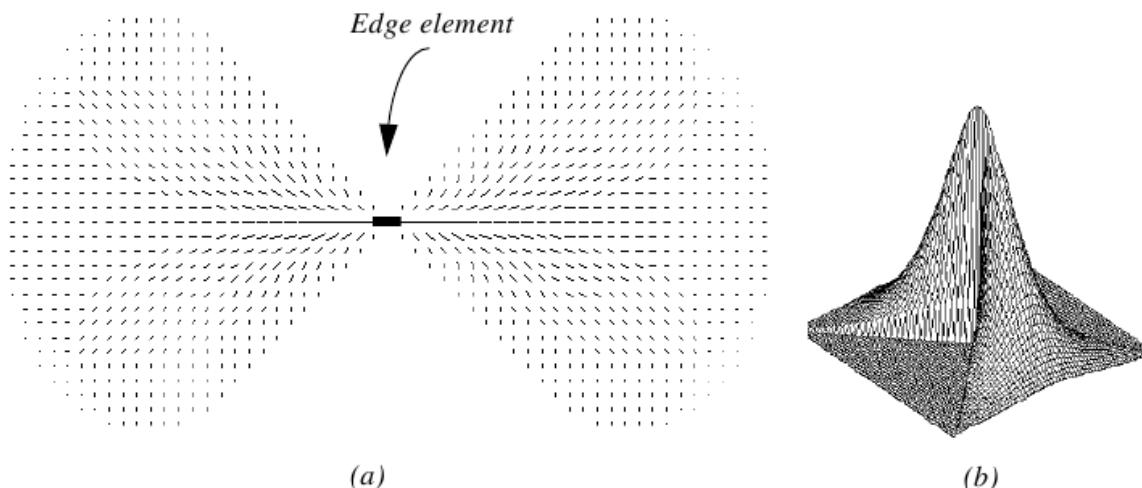
8. Z języka angielskiego – extension field.

Mimo, że w pracy autorzy nie skupiają się specjalnie na elementach, na których mają zamiar pracować – twierdzą, że może to być poziom piksli jak i cech, czyli krawędzi w obrazie – i raczej powinno się do tego podchodzić sceptycznie, to sama metoda może mieć sens, a obserwowane wyniki mogą napawać optymizmem.

Najprostszy opis metody polegałby na stwierdzeniu, że każdy element obrazu (pixel, bądź cecha – zależnie od poziomu abstrakcji) jest określony poza swoimi zwykłymi parametrami pozwalającymi na zidentyfikowanie jego w obrazie (na przykład położenie) przez parametry takie jak kąt i siła. Są one ściśle przypisane do każdego segmentu – na który cały obraz jest podzielony. Każdy segment zbiera informacje z innych segmentów, ze swojego otoczenia (zgodnie z konkretnymi wagami). Następnie liczona jest „zgodność wśród głosów” z punktu widzenia orientacji segmentów. Spośród tych segmentów, te z wysoką oceną są nazywane istotnymi⁹.

Pole rozszerzenia

Przedstawione poniżej – na rysunku 2.13 – pole rozszerzenia przedstawia zakres wpływu sąsiednich elementów krawędziowych na jeden konkretny. Co autorzy cytowanej pracy podkreślali istotny jest jedynie pozytywny wpływ, czyli nie ma możliwości wpływania negatywnego na wynik – siłę, jaka zostanie danemu elementowi krawędziowemu przypisana. Czego wprawdzie nie widać na załączonym rysunku, to fakt, że teoretycznie zakres wpływu na element krawędziowy jest nieograniczony.



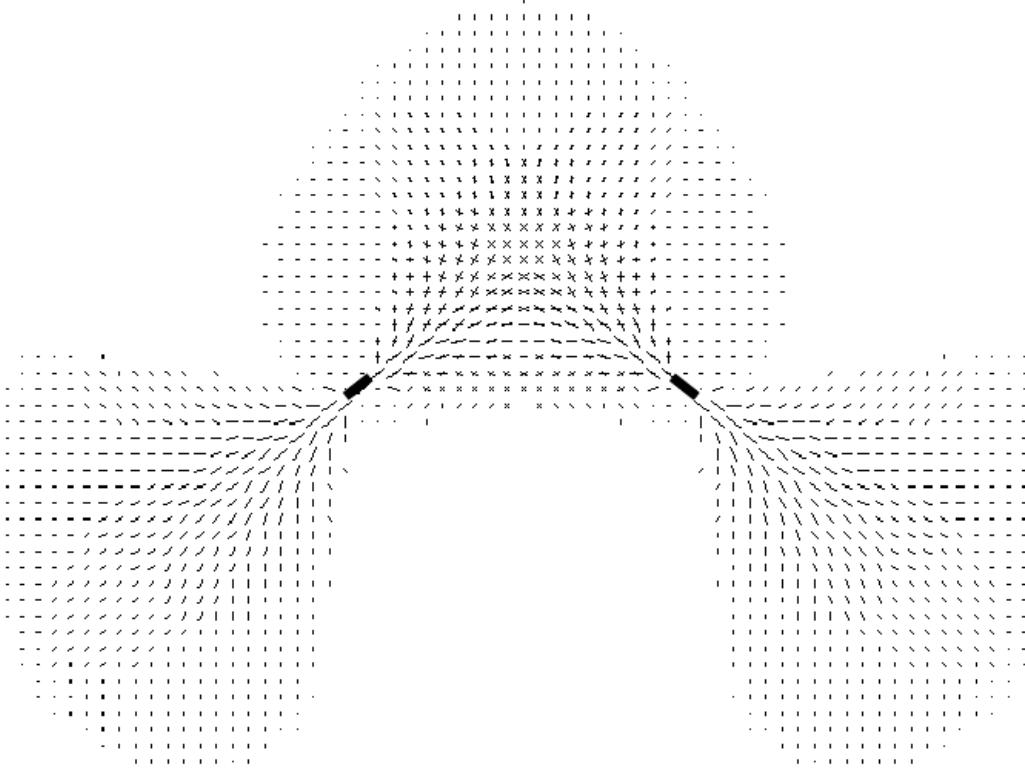
Rys. 2.13: Pole rozszerzenia opisujące wpływ sąsiednich segmentów na jeden konkretny ze względu na (a) kąt oraz (b) siłę [10].

Należy również przedstawić sytuacje, w których segmenty krawędziowe będą w pewnej istotnej¹⁰ odległości od siebie. Złączenie pól rozszerzenia pokaże wtedy jak jest interpretowany związek tych dwóch elementów krawędziowych i to w jaki sposób

9. Z języka angielskiego – salient.

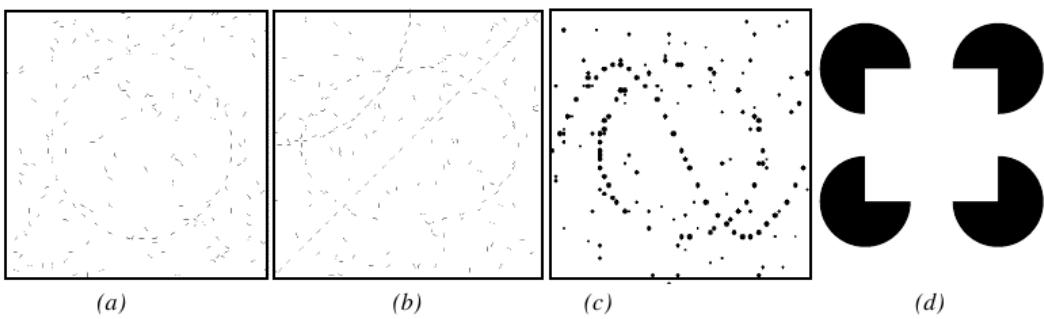
10. Takiej, że będzie ona miała znaczący wpływ na brany pod uwagę inny element krawędziowy.

będzie uogólniana część krzywej, która mogłaby je łączyć. Rysunek 2.14 przedstawia taką właśnie sytuację.



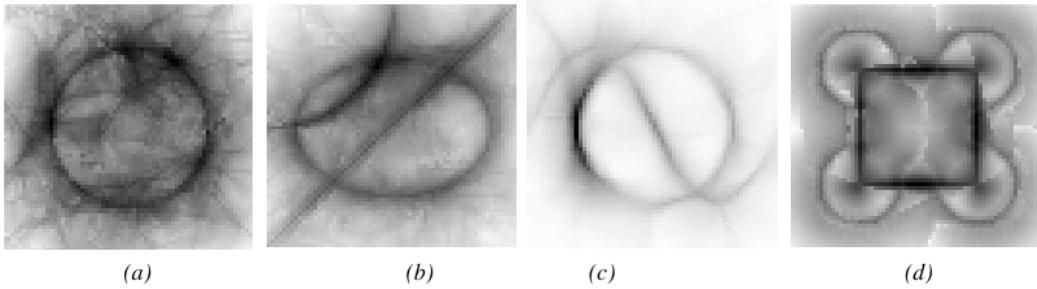
Rys. 2.14: Wzajemny wpływ dwóch pól rozszerzenia na siebie i krzywa która jest na tej podstawie wydobywana [10].

Przedstawiany model wprawdzie jest bardzo prosty – mimo, że nie przytacza się w tej pracy całej teorii matematycznej, która za tym stoi – ale wyniki, jakie można osiągnąć przetwarzając obrazy przedstawione w różnej formie są napawające optymizmem. Na rysunku 2.15 przedstawione są obrazy wejściowe (w formie krawędzi – a i b, kropek – c jak i obrazu – d) natomiast na rysunku 2.16 widać otrzymywane rezultaty, będące przedstawione w postaci samej mapy istotności.



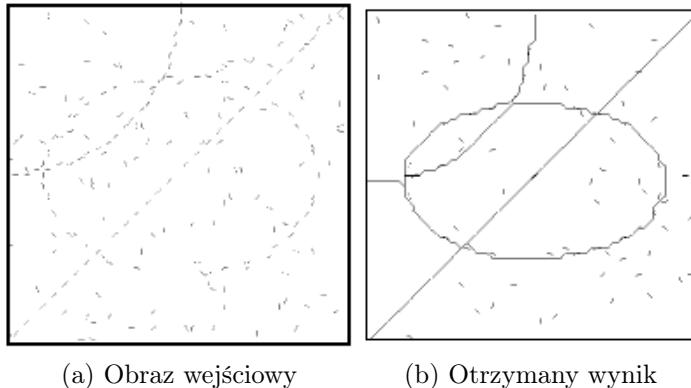
Rys. 2.15: Obrazy wejściowe dla modelu [10].

Po przetworzeniu tak przedstawionej mapy istotności do wyodrębnionych krawędzi można otrzymać obiecujące wyniki. Mimo, że w cytowanej pracy jest przedstawiony



Rys. 2.16: Otrzymywane wyniki przedstawione jako mapa istotności [10].

tylko jeden, to nic nie stoi na przeszkodzie, aby przypuszczać, że inne byłyby podobne. Ostateczny rezultat przedstawiony jest na rysunku 2.17. Odpowiada on obrazowi wejściowemu 2.15b oraz mapie istotności 2.16b. Zdaniem autora niniejszej pracy bardziej efektowne byłoby przedstawienie obrazu 2.15d po przetworzeniu przez model i pewnie zdanie czytelnika byłoby podobne, nie zmienia to jednak faktu, że w cytowanej pracy nie ma przedstawionych za pomocą krawędzi wyników modelu dla innych obrazów niż 2.15b. Odnosi się bowiem wrażenie, że model jedynie pozbył się kilku nieistotnych zakłóceń w obrazie wejściowym.



Rys. 2.17: Wynik pracy modelu 2.17b z obrazem 2.17a przedstawiony w formie krawędzi. [10].

2.2. Pełny model obliczeniowy kory wzrokowej człowieka

Przedstawiany w tej części model jest jednym z najczęściej spotykanych i cytowanych w czasopismach naukowych, jeśli kryterium będzie zbiór czterech słów kluczowych: object, recognition, visual, cortex. Jest nie tylko inspiracją dla prac magisterskich [17], doktoratów [22], ale i wniosków patentowych [25].

Praca autorów tego modelu jako jedna z nielicznych ma za zadanie nie tylko dobre symulowanie działania kory wzrokowej człowieka, ale i dobre odwzorowanie samej struktury kory wzrokowej człowieka. Model połączeń między elementami kory

wzrokowej człowieka, z jakiego wychodzą autorzy bardzo przypomina ten z cytowanej wcześniej pracy [28], z uwagi na to, że ich praca jest młodsza można wnioskować, że jest on aktualniejszy. Sam model, przedstawiany w tej pracy na rysunku 2.18 i zdaniem autora tej pracy bardzo dobrze obrazuje skomplikowanie, o którym wspominają naukowcy zajmujący się badaniem mózgu na codzień. Model ten nie jest specjalnie istotny z punktu widzenia zadania jakie ma wykonywać, o wiele ważniejsza jest struktura przedstawiająca jakby wniosek z tego modelu. Na rysunku 2.19 przedstawiona jest hierarchia połączeń między cechami, która jest zaimplementowana w modelu.

Sens takich połączeń, hierarchi można w bardzo prosty sposób wyjaśnić mając na uwadze to co się dzieje w ludzkim mózgu. Już na poziomie pierwszorzędowej kory wzrokowej, z informacji pobranych z ciała kolankowego bocznego pobierane, two-rzone są niewielkie krawędzie. Następnie – wyżej się będzie poruszać w hierarchii, z tym większymi obiektaymi ma się do czynienia. Niewielkie krawędzie łączą się bowiem w większe, a następnie w złożone cechy, reprezentujące grupy krawędzi, nie będące współliniowymi. Opisywane poziomy szczegółowości dobrze przedstawione są na rysunku 2.20 i należało by patrzeć równocześnie na rysunek 2.19, na którym każdy z poziomów szczegółowości jest umieszczony w hierarchii. Wyraźnie widać również, że poziom S4 jest ostatnim w którym występuje tak zwane nienadzorowane uczenie, a wyższe warstwy są już klasyfikatorami, przypisującymi konkretnym obiektom konkretne klasy. Co warto zauważyć, to fakt istnienia połączeń międzywarstwowych, który mimo, że po raz kolejny dobrze odzwierciedla połączenia w korze wzrokowej, komplikuje interpretację działania samego modelu.

Wchodząc natomiast w kwestie implementacyjne, czyli dotyczące realizacji modelu z rysunku 2.19 autorzy wniosku patentowego przedstawiają odpowiadające pierwszym czterem poziomom w hierarchii przekształcenia. Ich dokładny opis przedstawiony jest na rysunku 2.21. W literaturze nazywa się ten model słowem *HMAX*, które jak można przypuszczać wywodzi się od połączenia Hierarchiczności z operacją MAXimum, która jest jednym z najważniejszych przekształceń zachodzących w modelu.

Przedstawiony etap S1 na rysunku 2.21 stosuje opisane w załączniku A stosowanie filtra Gabora. Autorzy za istotne uznali nie tylko kąty, pod którymi funkcja Gabora jest w oknie zorientowana, ale i wielkość samego okna filtru. Co można przeczytać na opisywanym rysunku zdecydowali oni, że dużo istotniejsza jest skala od orientacji, przez co orientacje są tylko 4, natomiast skala 16. Każda orientacja na rysunku 2.21 oznaczona jest innym kolorem, a każde występujące pod sobą 4 różne kolory odpowiadają jednej skali – wielkości okna filtru Gabora.

W etapie C1 spośród otrzymanych wyników filtrowania jest wyciągane maximum ze względu na pozycję i skalę. W tej części wynikiem operacji maximum jest część przefiltrowanego obrazu, która wykazała się największą odpowiedzią spośród tych z sąsiedniej skali i orientacji.

Etap S2 przedstawia ocenę wszystkich otrzymanych w poprzednim etapie elementów obrazu w stosunku do zapamiętanych przez model prototypów – którymi są pewne

elementy opisujące krawędzie (przypis autora). Ocena przebiega poprzez policzenie odległości euklidesowej między elementem zapamiętanym w modelu, a fragmentem obrazu z poprzedniego etapu. Dodatkowo ocena ta jest argumentem pewnej funkcji Gauso-podobnej, aby dodatkowo zmniejszyć znaczenie fragmentów obrazu, którym dużo brakuje do danych elementów modelu.

Jak się można było spodziewać, etap C2 będzie polegał na zastosowaniu funkcji maximum globalnie, po wszystkich wynikach poprzedniego etapu.

Mimo, że przedstawiona – jedna z wielu istniejących – implementacja kończy przetwarzanie informacji wizualnych na poziomie S2, to można sobie wyobrazić kontynuację tego procesu, w którym na kolejnych poziomach są budowane coraz to bardziej złożone cechy. Na samym szczycie następuje natomiast klasyfikacja i przypisywanie otrzymanego w ten sposób wyniku do konkretnej klasy.

2.2.1. Rozpoznawanie obiektów wywodzące się z podziału na komponenty

Co jest podkreślone bardzo często w świecie publikacji, przedstawiony model jest ciężki obliczeniowo i istnieje bardzo wiele jego modyfikacji, które nastawione są bardzo często na zwiększenie jego wydajności. Podobnie planował to zrobić autor wymienianej na początku tej części pracy magisterskiej [17].

Kluczem do zyskania na wydajności – zdaniem niektórych naukowców – jest rozpoznawanie obrazów ze względu na wydzielone w nich komponenty. W ten sposób nie będzie trzeba za każdym razem wielokrotnie przetwarzać całego obrazu, lecz – w zależności od implementacji – jego elementy, przez co duża część obliczeń może zostać zredukowana. Rzecz jasna nie ma nic za darmo i zysk z realizacji części obliczeń jest równoważny utracie ogólności w rozpoznawaniu obiektów na obrazie. Ponieważ po ekstrakcji komponentów z obrazu, może i stworzenia sieci komponentów połączonych pewnymi relacjami między sobą będzie można przeprowadzić jedynie operację sprawdzenia, czy w zadany obrazie można rozpoznać cechy, jakie przez model zostały wyszczególnione. Bardzo dobrze to ograniczenie, jak i sam proces tworzenia modelu obrazuje rysunek 2.22.

Wyraźnie widać, że na początku należy wyszukać w obrazie punktów szczególnych – co niestety jest trudne do zdefiniowania. Drugim etapem jest wyznaczenie pewnego otoczenia tych punktów szczególnych oraz połączenia tych otoczeń relacjami na przykład odległości.

2.3. Podsumowanie

W dalszej części tej pracy przedstawiony będzie szczegółowy opis proponowanego przez autora podejścia, które w dużej mierze czerpać będzie z przedstawionych w literaturze modeli mory wzrokowej człowieka. Wstępnie można by nakreślić kroki, jakimi warto by było zrealizować cel postawiony w temacie tej pracy.

I Odzyskiwanie krawędzi z obrazów – czyli zastosowanie filtrów Gabora do wyszczególnienia krawędzi w zadanych obrazach.

Zawartość załącznika A świadczy o realizacji tego etapu.

II Ekstrakcja krawędzi z przefiltrowanych obrazów – wydobycie krawędzi czyli pewnych kresek – mających reprezentować krawędzie obiektów występujących w niewielkich fragmentach obrazu.

Nawiązując do wcześniejszych części tej pracy – jest to etap trudny i przy realizacji pod uwagę będzie branych wiele różnych algorytmów, wstępnie jednak nieokreślonych.

III Integracja konturów – pozostawienie w obrazie przedstawionym za pomocą krawędzi jedynie tych, które są istotne ze względu na to, że są częścią złożonej krawędzi, a nie zakłóceniem.

Wstępnie można by założyć wykorzystanie jednego z opisanych wcześniej podejść 2.1.4, 2.1.5 lub 2.1.6, ale autor wstępnie nie chce się przywiązywać specjalnie do konkretnego rozwiązania.

IV Budowanie cech złożonych – składanie krawędzi z poprzedniego etapu w większe struktury.

Podobnie jak to jest zaproponowane w części 2.2, liczba – wprawdzie nie cytowanych, ale istniejących – prac wskazuje na to, że hierarchiczne podejście jest dobrym rozwiązaniem.

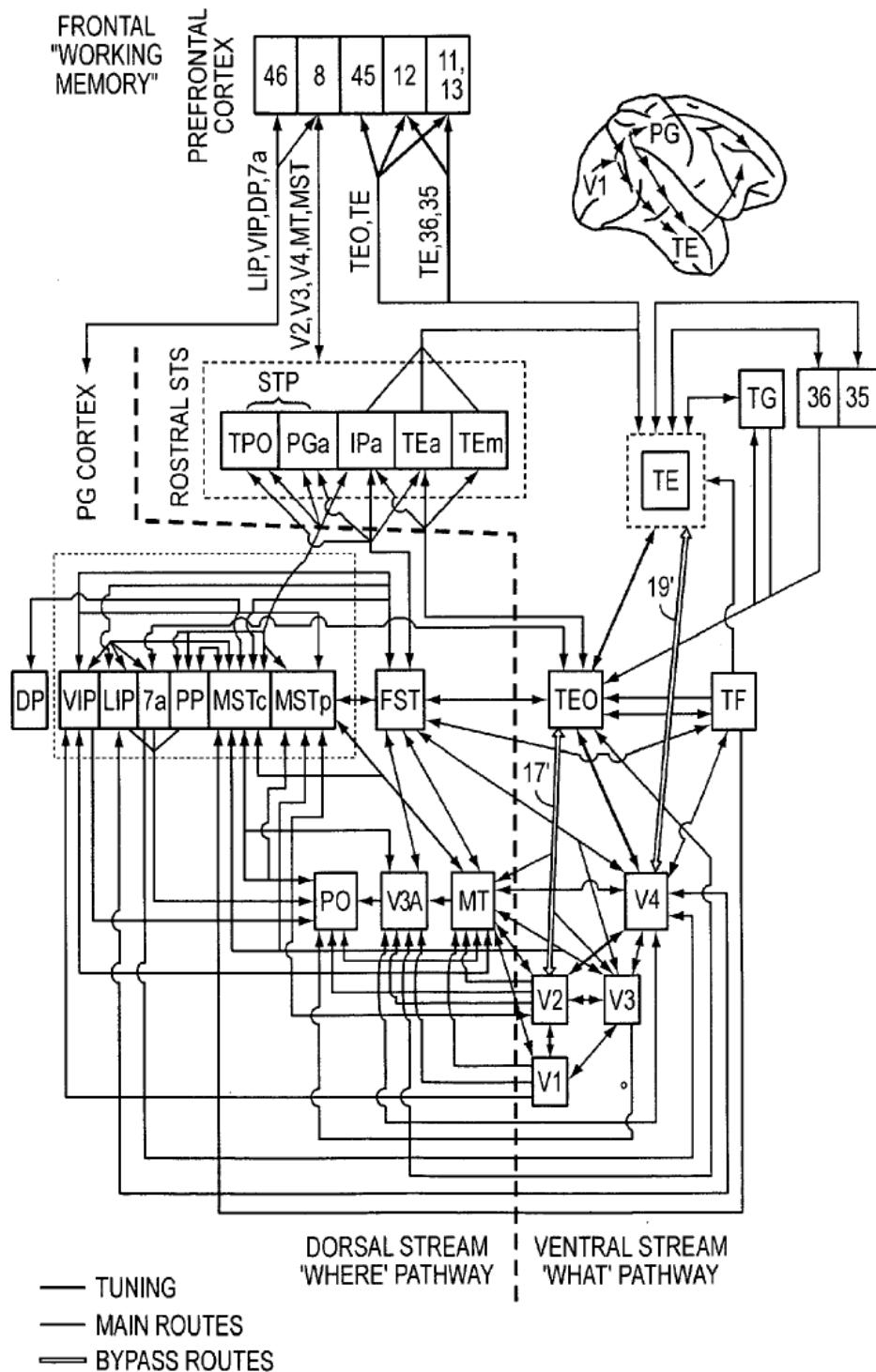
V Podział na obiekty występujące w obrazie – podział stworzonych w poprzednim etapie szkieletów – reprezentujących odnalezione w obrazie krawędzie – na obiekty ze względu na różnicę w kolorach.

Coś, co wydaje się być oczywiste. Przegląd literatury jednak nie wskazuje, żeby było to popularne rozwiązanie.

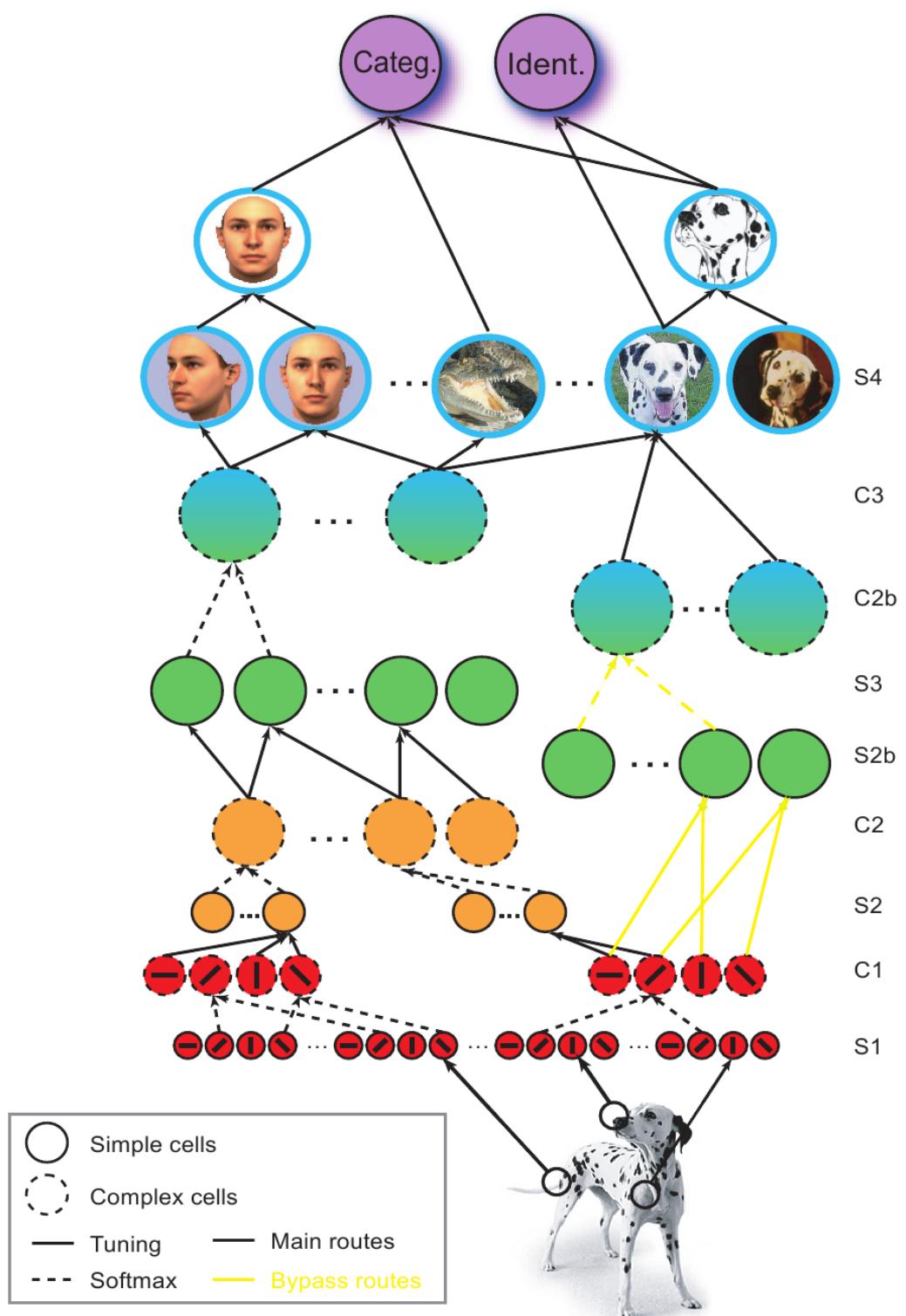
VI Klasyfikacja wydzielonych obiektów – przypisanie odnalezionemu obiekowi jego klasy – nazwanie obiektu z obrazu.

W niektórych pracach jest ten etap rozwiązywany jako sieć neuronowa – która z definicji jest klasyfikatorem. Autor rozważa zastosowanie w tym etapie sieci neuronowej, natomiast planuje również zbudowanie ontologii z klas, do których będą należeć obrazy dla stworzonego modelu. Konieczne jest również zebranie odpowiedniej liczby samych obrazów, na których model będzie pracował, ucząc się i szukając podobnych – czyli – realizując temat. Konieczna jest dodatkowo różnorodność obrazów, aby ontologia mogła prezentować pewną hierarchię.

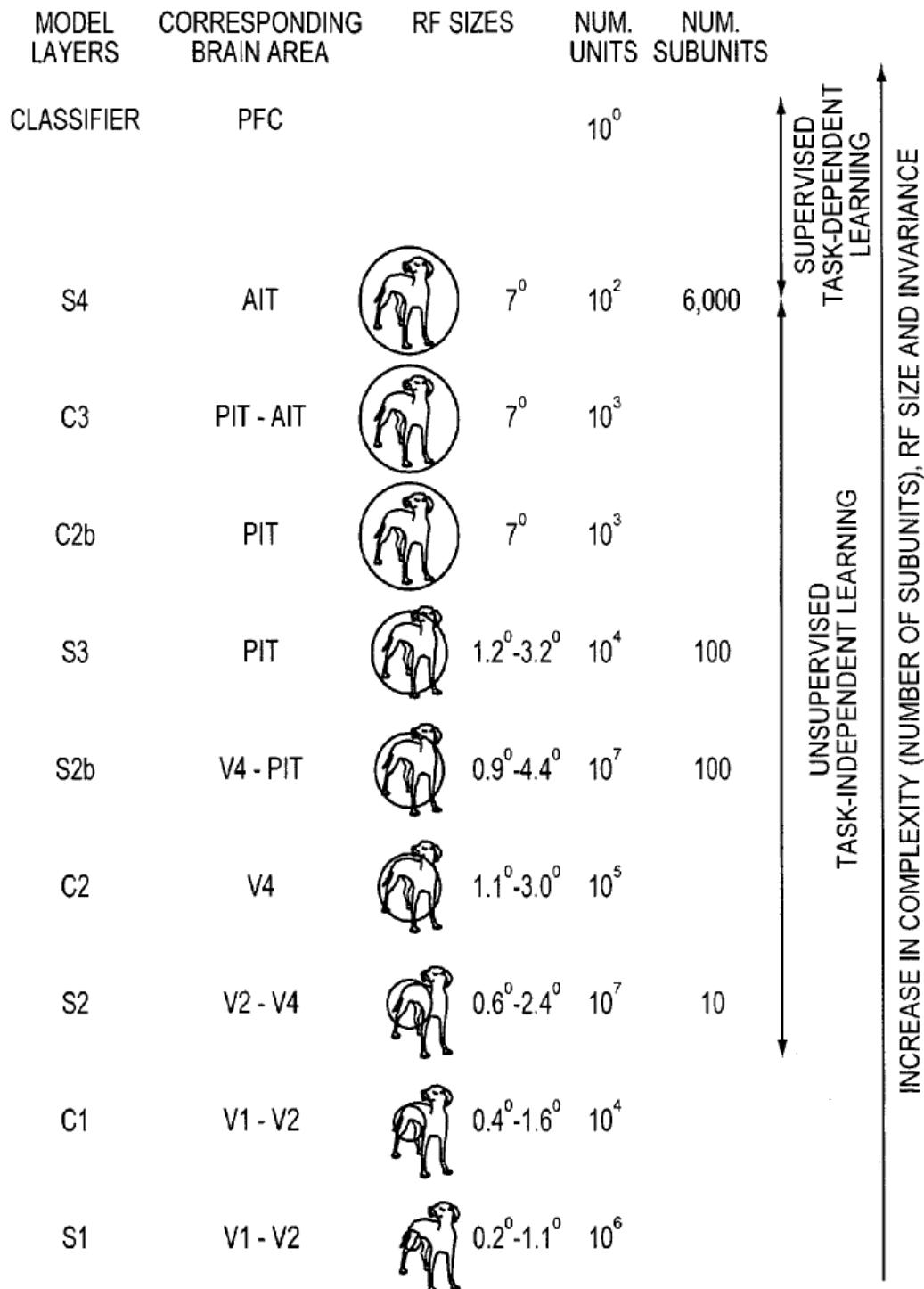
Autor planuje realizację tych etapów we właściwie takiej kolejności jak na liście powyżej, z zastrzeżeniem jednak, że zbieranie obrazów oraz budowanie ontologii z ostatniego etapu, ponieważ nie zależy od żadnego z nich, może być przeprowadzone równolegle z dowolnym.



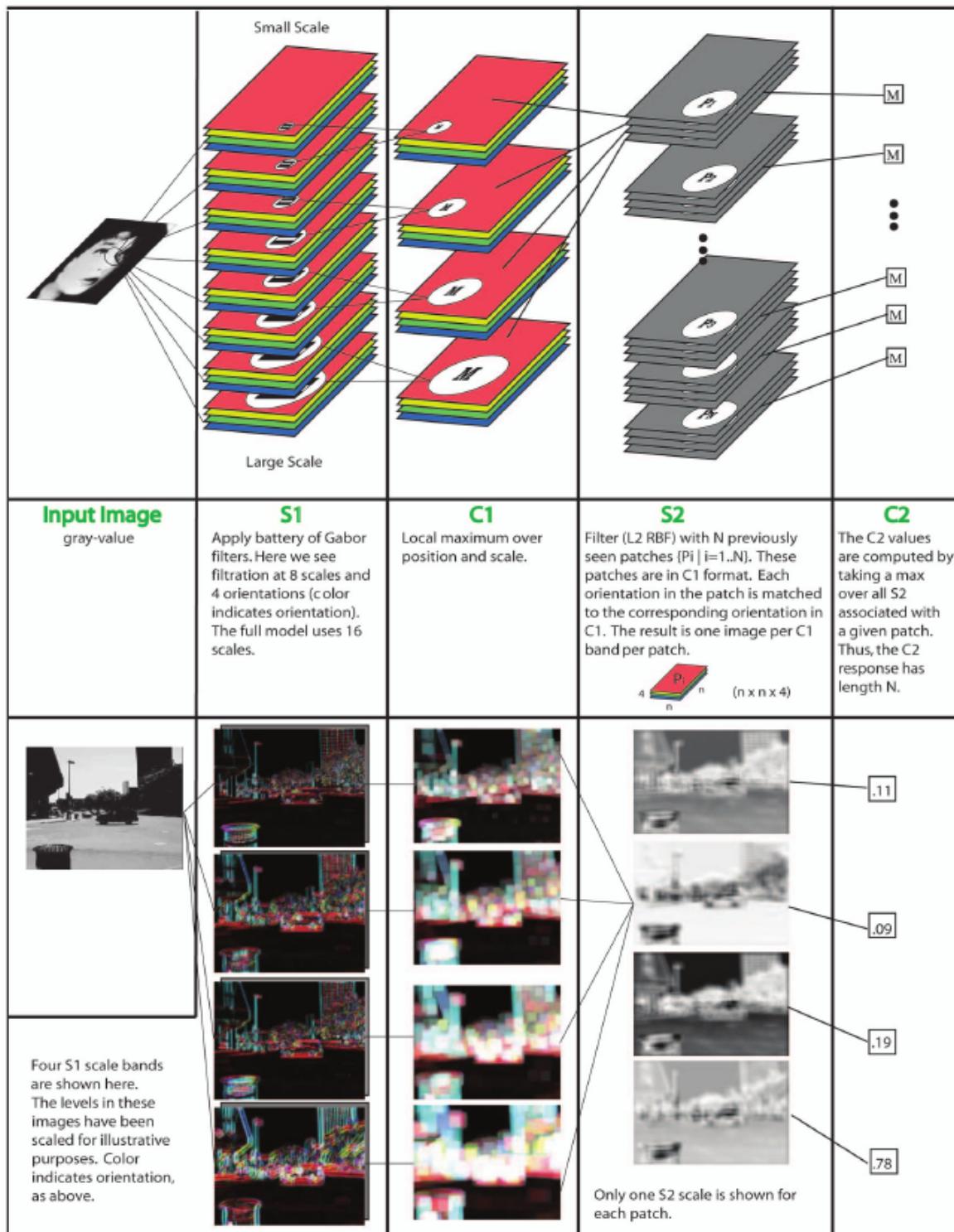
Rys. 2.18: Model przedstawiający połączenia między elementami kory wzrokowej człowieka oraz jego podział na dwa szlaki: brzuszny i grzbietowy, które ściśle ze sobą są powiązane [25].



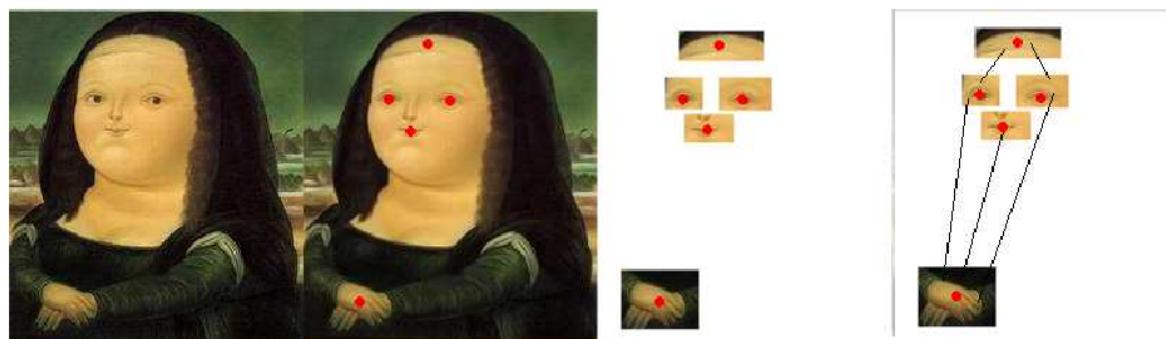
Rys. 2.19: Struktura odzwierciedlająca hierarchię połączeń w modelu rozpoznawania obrazów [24].



Rys. 2.20: Odzwierciedlenie zakresu szczegółowości na kolejnych poziomach modelu przedstawionego na rysunku 2.19 [25].



Rys. 2.21: Przedstawienie szczegółów dotyczących podstawowych bloków, z których jest złożony model z rysunku 2.19 [26].



Rys. 2.22: Kolejne kroki tworzenia modelu komponentów [17].

Dodatek A

Filtr Gabora

A.1. Teoria

Filtr w grafice komputerowej jest tak zwaną konwolucją dyskretną [27] obrazu. Dla przypadku 2 wymiarów będzie ona przedstawiona poniżej.

$$L'(m, n) = (w \times L)(m, n) = \sum_{i,j \in K} L(m - i, n - j)w(i, j)$$

Należy wyjaśnić, że w powyższym przykładzie L jest obrazem początkowym, gdzie $L(x, y)$ jest punktem w tym obrazie określonym współrzędnymi x i y . Natomiast $w(x, y)$ jest filtrem, którego składowe x i y są tak samo interpretowane. Wynikiem przedstawionej w ten sposób filtracji obrazu jest L' .

Można powiedzieć, że przy filtrowaniu posiadamy 3 macierze/tablice(2D) różnych rozmiarów. Z czego przyjmuje się, że macierze L i L' są tych samych rozmiarów, a macierz w jest innych rozmiarów (najlepiej aby jej rozmiary przedstawiane były za pomocą 2 liczb nieparzystych) i jest dużo mniejsza od L . Macierz w jest „oknem filtra”.

Filtrowanie jest procesem w którym dla każdego piksela z obrazu wejściowego – nawiązując do oznaczeń zaproponowanych powyżej – L kolor odpowiadającego mu piksla w obrazie wyjściowym (L') jest zależny od ważonej sumy wartości koloru piksli go otaczających. Wagi dla tego przekształcenia są określone przez okno filtra i przyjmuje się, że środkowemu punktowi okna filtra odpowiada wybrany wcześniej piksel w obrazie L . Dodatkowo wymaga się, aby zostały określone warunki na podstawie których będzie można to przekształcenie stosować na krawędziach obrazu, dochodzi tam bowiem do sytuacji, w których okno filtra wykracza poza obraz. Przyjmuje się, że w takich wypadkach nieistniejące piksle nie biorą udziału w wyliczeniach wartości koloru dla danego piksła leżącego przy krawędzi. Należy jednak dodatkowo zwrócić uwagę na fakt, że obraz jest dwuwymiarową tablicą kolorów, z których każdy jest w określonych granicach. Wymagane jest więc, aby po przekształceniu obrazu danym filtrem doprowadzić jego wartości kolorów w taki sposób, aby mógł on odzwierciedlać

obraz.

Przykładowo obraz może być dwuwymiarową tablicą piksli, z których każdy jest określonego koloru, a kolor ten jest przedstawiony np. jako wartość jasności (dla obrazów czarno-białych) i mieści się w zakresie (0, 255). Po przekształceniu filtrem o oknie wielkości 3×3 , w którym wartości są z zakresu $(-1, 1)$ należałyby wartości w obrazie wynikowym tak przemnożyć, aby również przedstawiały określony wcześniej zakres.

A.1.1. Funkcja Gabora

Funkcja Gabora [18], która jest głównym składnikiem filtru Gabora jest bardzo prostym złożeniem dwóch, tak naprawdę elementarnych funkcji. Jako złożenie funkcji autor ma na myśli prosty iloczyn – jak to jest przedstawione we wzorze A.1.

$$g(x, y) = s(x, y)w_r(x, y) \quad (\text{A.1})$$

W przedstawionym powyżej wzorze należy wyszczególnić dwa elementy:

$s(x, y)$ – jest zespoloną sinusoidą, nazywaną nośnikiem¹

$w_r(x, y)$ – jest dwuwymiarową funkcją o kształcie funkcji Gaussa, nazywaną powłoką²

Nośnik

Zespolona sinusoida jest odwzorowaniem funkcji A.2 na płaszczyźnie.

$$s(x, y) = \exp(j(2\pi(u_0x + v_0y) + P)) \quad (\text{A.2})$$

W podanym wyżej wzorze można wydzielić parametry:

P – faza sinusoidy

(u_0, v_0) – wektor, który opisuje częstotliwość przestrzenną sinusoidy³

Modelowanie funkcji Gabora bardzo wygodnie jest przedstawiać jako złożenie części urojonej i rzeczywistej, tak jak to jest zobrazowane we wzorach A.3 i A.4.

$$\Re(s(x, y)) = \cos(2\pi(u_0x + v_0y) + P) \quad (\text{A.3})$$

$$\Im(s(x, y)) = \sin(2\pi(u_0x + v_0y) + P) \quad (\text{A.4})$$

Chcąc przekształcić tak przedstawioną funkcję do współrzędnych biegunowych – co w efekcie dużo uprości – wymagane jest zastosowanie kilku przekształceń, które bezpośrednio wynikają z zastosowania kilku reguł matematycznych.

$$F_0 = \sqrt{u_0^2 + v_0^2} \quad (\text{A.5})$$

$$\omega_0 = \tan^{-1}\left(\frac{v_0}{u_0}\right) \quad (\text{A.6})$$

Gdzie:

F_0 – wzmocnienie⁴

1. Z języka angielskiego carrier.
2. Z języka angielskiego envelope.
3. Z języka angielskiego spatial frequency.
4. Z języka angielskiego magnitude.

ω_0 – kierunek

Rzecz jasna można odwrócić sytuację i przedstawić u_0 i v_0 jako funkcje F_0 i ω_0 , jak to jest przedstawione we wzorach A.7 i A.8.

$$u_0 = F_0 \cos \omega_0 \quad (\text{A.7})$$

$$v_0 = F_0 \sin \omega_0 \quad (\text{A.8})$$

W tej reprezentacji funkcja ma następującą postać.

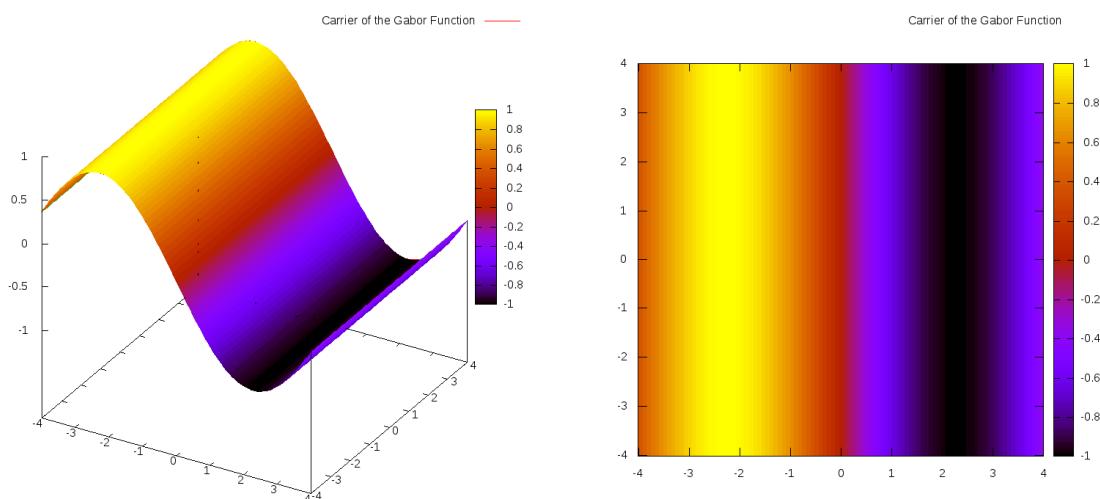
$$s(x, y) = \exp(j(2\pi F_0(x \cos \omega_0 + y \sin \omega_0) + P)) \quad (\text{A.9})$$

Co podczas implementacji najprościej przedstawić jako następującą funkcję.

$$s(x, y) = \sin(2\pi F_0(x \cos \omega_0 + y \sin \omega_0) + P) \quad (\text{A.10})$$

Takie przedstawienie nośnika czyni z niego jakby funkcję sinus w przestrzeni, która jest rozciągnięta w tym trzecim wymiarze. Trochę przypomina płaską falę na wodzie, co obrazuje dobrze rysunek A.1.

Sterując teraz wartością ω_0 można obracać funkcję wokół punktu $(0, 0)$.



Rys. A.1: Nośnik, element funkcji Gabora, przedstawiony w przestrzeni i jako mapa.

Powłoka

Powłoka, która jest dwójwymiarowym odpowiednikiem funkcji Gausa ma postać, przedstawioną we wzorze A.11.

$$W_r(x, y) = K \exp(-\pi(a^2(x - x_0)^2_r + b^2(y - y_0)^2_r)) \quad (\text{A.11})$$

Gdzie:

(x_0, y_0) – to szczyt funkcji

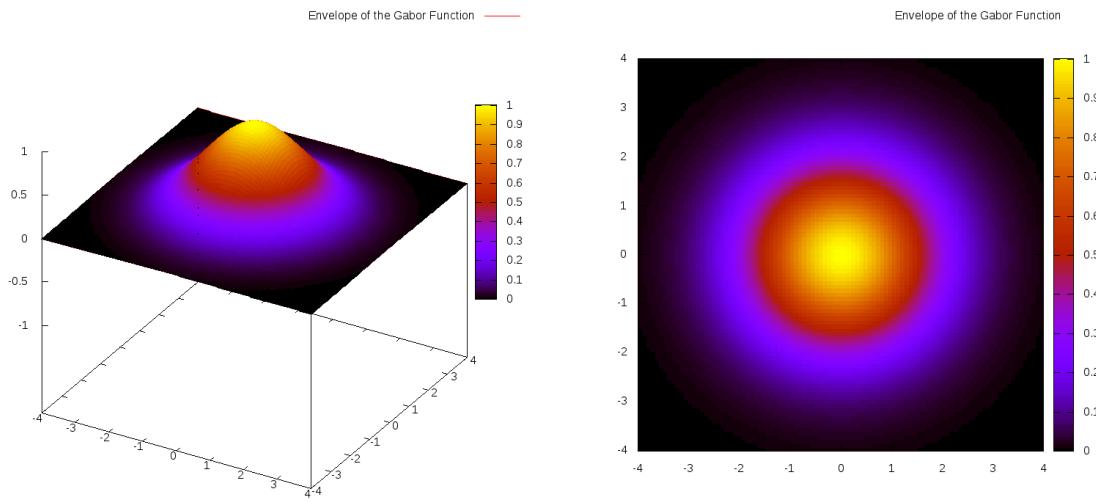
a i b – parametry odpowiedzialne za skalę powłoki w funkcji Gabora

Przedstawiony w powyższym wzorze A.11 indeks dolny r odpowiada za rotację całej powłoki. Odpowiadające nawiasom z indeksem r wzory przedstawione są poniżej.

$$(x - x_0)_r = (x - x_0) \cos \theta + (y - y_0) \sin \theta \quad (\text{A.12})$$

$$(y - y_0)_r = -(x - x_0) \sin \theta + (y - y_0) \cos \theta \quad (\text{A.13})$$

W efekcie czego można uzyskać powłokę przedstawioną na rysunku A.2.



Rys. A.2: Powłoka, element funkcji Gabora, przedstawiony w przestrzeni i jako mapa.

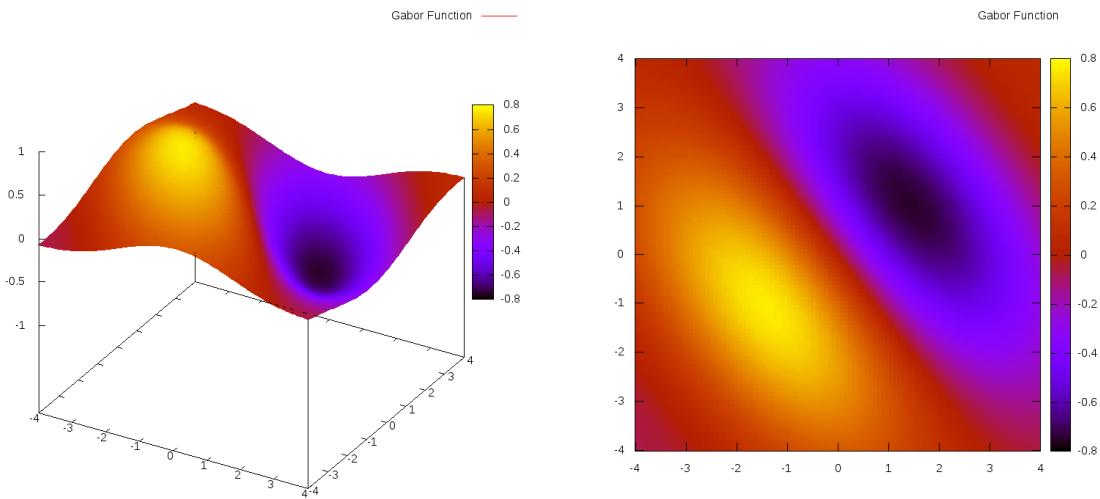
Złożenie nośnika i powłoki

Składając opisane wyżej funkcje ze sobą – tak jak to jest przedstawione we wzorze A.1 – można otrzymać funkcje podobne do tej, przedstawionej na rysunku A.3.

A.1.2. Filtr Gabora

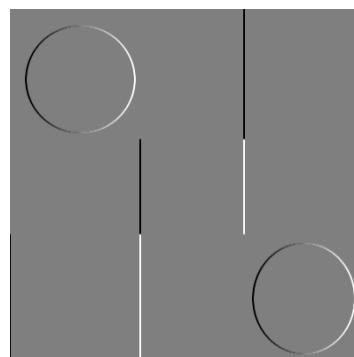
Filtr Gabora jest połączeniem dwóch wymienionych już wcześniej technik, z taką różnicą, że jako okno filtru będzie w nim występować odpowiednio przeskalowana funkcja Gabora. Zgodnie z przedstawionym na samym początku opisem można wyprowadzić algorytm, za pomocą którego będzie można wykonać filtrowanie obrazu. Proste filtrowanie obrazu przedstawione jest w algorytmie 1.

Filtrowanie przedstawione w algorytmie 1 pokazuje samą zasadę działania wszystkich filtrów w grafice komputerowej. Algorytm zakłada, że macierz filtru – określana na początku załącznika jako okno filtru – jest prostokątem, którego boki są dowolną liczbą naturalną nieparzystą. Jest to warunek, w który powoduje, że łatwiej jest sobie wyobrazić sposób działania filtru i jego sens. Dodatkowo należy wspomnieć, że zaprezentowany algorytm nie uwzględnia sytuacji w których dla piksla będącego na



Rys. A.3: Funkcja Gabora przedstawiona w przestrzeni i jako mapa.

krawędzi w obrazie źródłowym, wyliczana jest suma ważona wraz z pikslami, które są poza obrazem. Rzecz jasna jest to sytuacja niedopuszczalna i problem należałoby rozwiązać sprawdzając, w kroku 11 czy piksel z otoczenia istnieje w obrazie źródłowym. Ostatnia modyfikacja, jaka powinna zostać przeprowadzona na algorytmie polega na zachowaniu zakresu wartości kolorów w pikslach. Wymagane jest choćby zastosowanie transformacji z ostatniego akapitu części A.1.



Rys. A.4: Wynik filtrowania obrazu A.5a pierwotną wersją filtru Gabora.

Dodatkowo można zauważać pewną prawidłowość podczas filtrowania obrazów, która jest istotna ze względów stricte wydajnościowych. O ile bowiem samego procesu filtrowania pojedynczego obrazu nie można przyspieszyć, to filtrowanie obrazu mające na celu wyodrębnienie krawędzi filtrem Gabora – przedstawiane w obrazach A.5b-k – można dwukrotnie przyspieszyć. Należy bowiem zwrócić uwagę na to, że filtrując obraz (choćby A.5a) filtr będzie w wynikowym obrazie ciemniejsze kolory dawać krawędziom występującym w okolicach piksli, których suma ważona będzie największa. Patrząc na krok 11 algorytmu 1 widać, że wagi do tej sumy są umieszczone w oknie filtru, przez co

algorytm 1 Zastosowanie dwuwymiarowego filtru w przetwarzaniu obrazów.

```
1: picture  $\leftarrow$  obraz wejściowy
2: output  $\leftarrow$  obraz wyjściowy
3: filterMatrix  $\leftarrow$  filtr
4:  $a \leftarrow \lceil filterMatrix.width/2 \rceil$ 
5:  $b \leftarrow \lceil filterMatrix.height/2 \rceil$ 
6: for  $x = 1 \rightarrow picture.width$  do
7:   for  $y = 1 \rightarrow picture.height$  do
8:     value  $\leftarrow 0$ 
9:     for  $i = 1 \rightarrow filterMatrix.width$  do
10:      for  $j = 1 \rightarrow filterMatrix.height$  do
11:        value  $\leftarrow value + picture[x - a + i][y - b + j] \cdot filterMatrix[i][j]$ 
12:      end for
13:    end for
14:    output[ $x$ ][ $y$ ]  $\leftarrow$  value
15:  end for
16: end for
```

maksymalna wartość będzie przypadać na piksel, w którego otoczeniu piksle będą tak ułożone, że minusy przy iloczynie się skrócają. I w pierwotnej wersji algorytmu jedynie o to chodziło. Zwracając jednak uwagę na zaistniały fakt można by powiedzieć, że odwrotna sytuacja zostanie zaobserwowana gdy filtrowaniu obrazu poddany zostanie fragment obrazu, w którym krawędź jest odwrotna, w sensie minusy przy iloczynie się nie skrócają a piksel otrzyma minimalną wartość. Obrazek przedstawiający zaistniałą sytuację przedstawiony jest na rysunku A.4. Aby takie sytuacje wyeliminować, można zastosować niewielkie uproszczenie kodu z algorytmu 1 w którym w kroku 14 będzie wykonywać dodatkowe sprawdzenie, takie jak to, przedstawione w algorytmie 2. Dzięki czemu będzie można otrzymywać rezultaty podobne do tych z rysunku A.5.

algorytm 2 Modyfikacja kroku 14 algorytmu 1.

```
1: if value  $< 0$  then
2:   output[ $x$ ][ $y$ ]  $\leftarrow -value$ 
3: else
4:   output[ $x$ ][ $y$ ]  $\leftarrow value$ 
5: end if
```

A.1.3. Przykład filtrowania

Filtrowanie z zastosowaniem zmodyfikowanego algorytmu 1 na obrazie monochromatycznym daje ciekawe rezultaty, jak pokazuje rysunek A.5.

Co warto nadmienić, filtr Gabora przy jednokrotnym zastosowaniu nie wykryje wszystkich krawędzi w obrazie. Jak można zobaczyć na obrazie A.5b-k przy konkretnej wartością kąta ω_0 wyraźniejsze są konkretne krawędzie. Istotne jest więc wyodrębnienie krawędzi pod wszystkimi kątami (z dowolną dokładnością) i połączenie obrazów w taki sposób, że dla każdego konkretnego piksla w obrazie brane będzie maksymalne

nasycenie koloru z wszystkich obrazów składowych.

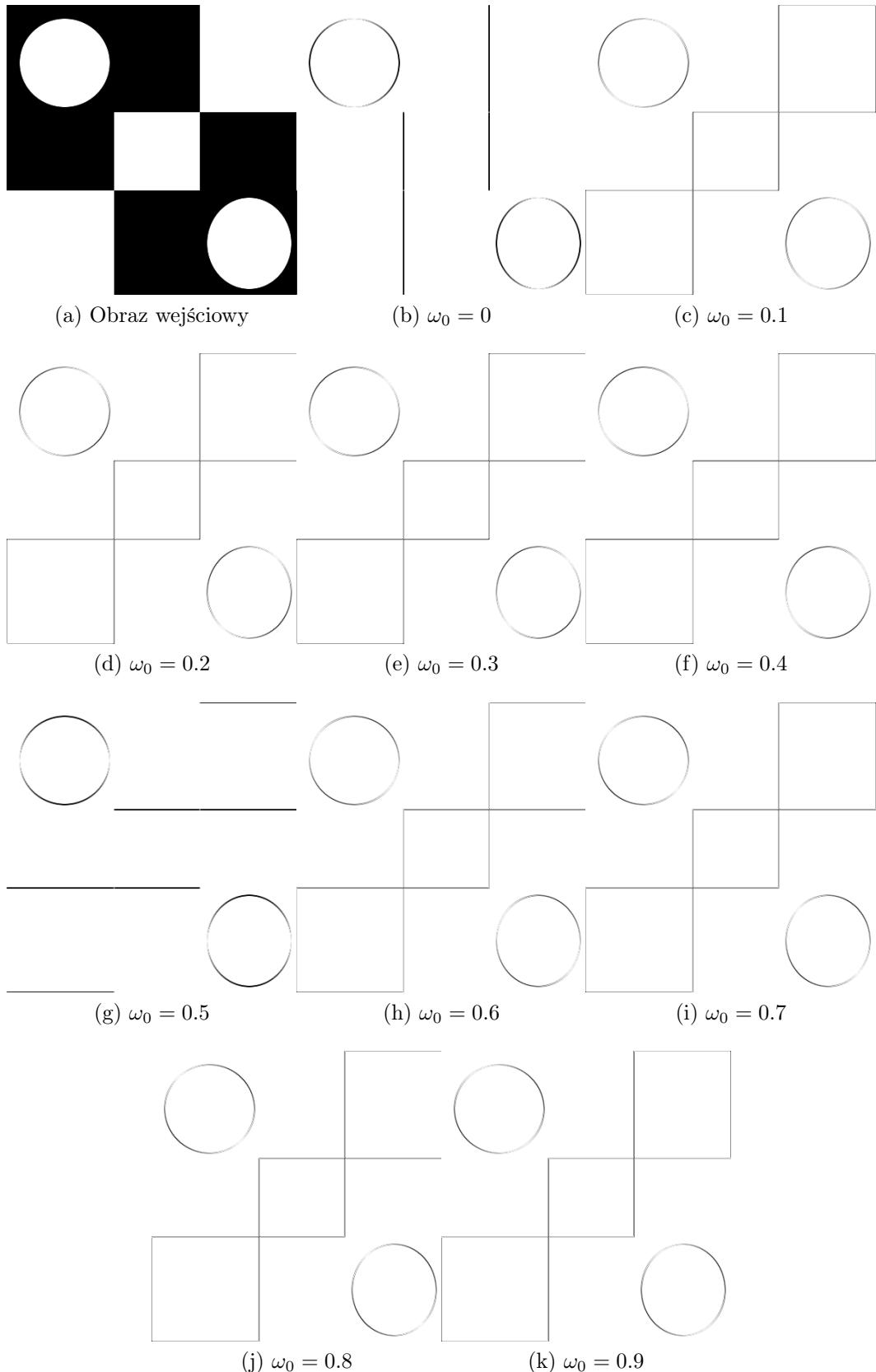
Rysunek A.6 przedstawia obraz wejściowy i wynik etapu filtrowania, na którym bardzo dobrze są widoczne krawędzie obrazu.

A.2. Dyskusja i wnioski

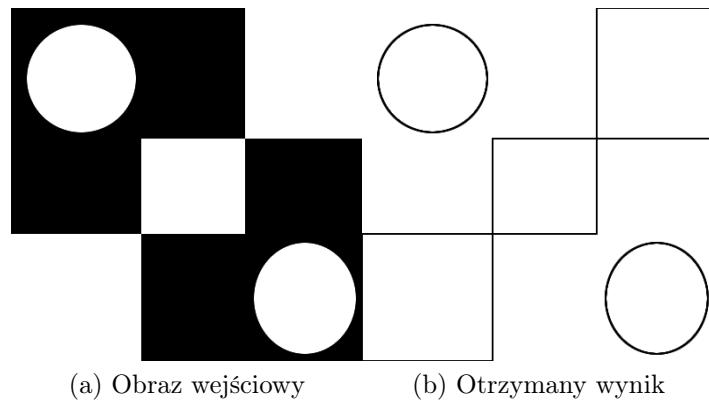
Wyniki pracy zaprezentowane w poprzednich etapach świadczą o skuteczności znaczania krawędzi w obrazach. Co było już podkreślone w pierwszych częściach tego dokumentu – nie tylko Gabor wpadł na dobry pomysł, natura sama od „zawsze” stosuje coś podobnego w korze wzrokowej [13]. Ale co należy podkreślić, sam filtr Gabora nie jest skutecznym rozwiążaniem. Przedstawione poniżej, na rysunku A.7 obrazki a-f pokazują, że istotny jest preprocessing⁵ obrazów, aby zmniejszyć liczbę występujących w obrazie zakłóceń. Dodatkowo widać, że nie tylko kąt funkcji sinus – będącej składową filtru Gabora – jest istotny, ale również wielkość okna filtru. Mimo, że obrazy są tendencyjnie pod względem prezentacji szczegółowych obiektów, na których widać wyraźne krawędzie, można sobie wyobrazić jednak obrazy z mniej wyraźnymi krawędziami, na których większe okno lepiej by sobie radziło.

Co widać teraz wyraźnie na przedstawianych obrazach, to fakt, że wyraźne oprócz środkowych części obrazu są też krawędzie. Dla nich bowiem proces filtrowania nie mógł być przeprowadzony z uwagi na brak sąsiednich piksli, które były by brane spoza obrazu. Przy dalszym przetwarzaniu obrazów wypadałoby tych krawędzi nie brać pod uwagę.

5. Zapożyczone z języka angielskiego – wstępne przetworzenie.



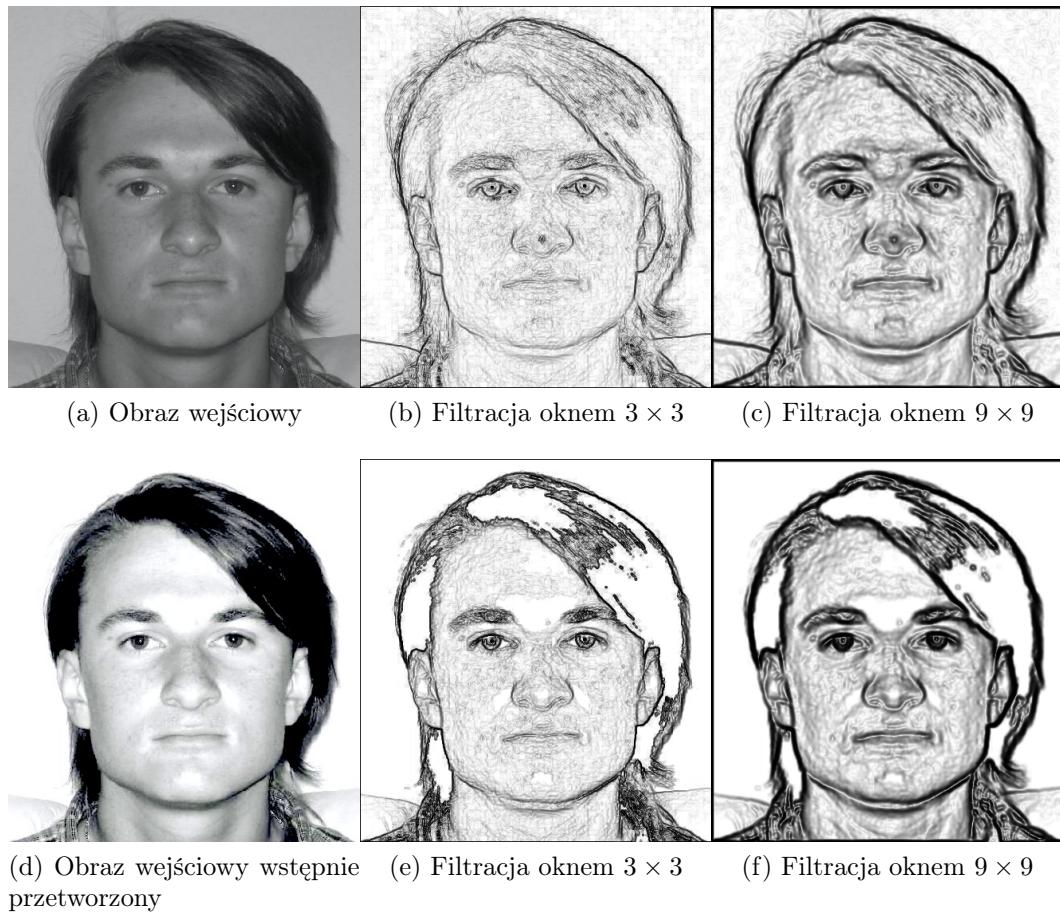
Rys. A.5: Wyniki filtrowania obrazu A.5a filtrami Gabora w których ω_0 jest w przedziale $(0, 1)$.



(a) Obraz wejściowy

(b) Otrzymany wynik

Rys. A.6: Wyniki filtrowania obrazu A.6a filtrami Gabora oraz operacji połączenia otrzymywanych rezultatów w jeden obraz.



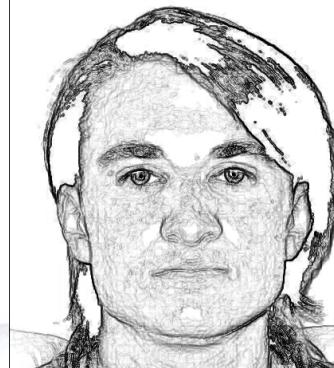
(a) Obraz wejściowy

(b) Filtracja oknem 3×3

(c) Filtracja oknem 9×9



(d) Obraz wejściowy wstępnie
przetworzony



(e) Filtracja oknem 3×3



(f) Filtracja oknem 9×9

Rys. A.7: Wyniki filtrowania obrazów A.7a i A.7d filtrami Gabora w których różna była wielkość okna filtru.

Spis rysunków

1.1	Punkty A i B przedstawiają pola wzbudzeń, neuronów z ciała kolankowatego bocznego, co również odpowiada komórkom siatkówki oka. Punkty C-G przedstawiają pola wzbudzeń neuronów pierwotnej kory wzrokowej, które są wrażliwe na istniejące w obrazie kreski, bądź krawędzie pod określonym kątem, lecz różnych rozmiarów. Można sobie wyobrazić, że w korze wzrokowej dla bardzo wielu nachyleń kąta są podobne grupy neuronów [12].	6
2.1	Przykład działania algorytmu DICONEX na obrazie w skali szarości, o rozmiarach 6×6 [23].	11
2.2	Wynik (b) działania algorytmu DICONEX na obrazie niewielkiej rozdzielczości (a) [23].	12
2.3	Model tablicy cech wizualnych, prostokąty symbolizują struktury danych, elipsy natomiast transformacje [20].	13
2.4	Krawędzie z modelu tablicy cech wizualnych, mające reprezentować krawędzie wyodrębnione w obrazie [21].	14
2.5	Rezultaty pracy modelu z obrazem: a) obraz wejściowy, b) wyjście z modelu [21].	14
2.6	Model hiperkolumny, zaproponowany przez autorów pracy [3], powstały na podstawie badań kory wzrokowej kota.	15
2.7	Przestrzeń hiperkolumn w modelu integracji konturów [16].	16
2.8	Architektura częściowa sieci prezentującej podstawowe elementy składowe modelu z pracy [16].	17
2.9	Pole mające wpływ na konkretny jeden segment, w zależności od funkcji odległości $J_{i\theta,j\theta'}$ oraz $W_{i\theta,j\theta'}$ [16].	18
2.10	Wyniki działania opisanego wcześniej algorytmu. Praca przedstawia zdjęcie najczęściej poddawane obróbce cyfrowej – Lenę. Wyraźnie widać, że krawędzie zostały odzyskane w praktycznie doskonały sposób [16].	19
2.11	Przedstawienie porównania modeli integracji konturów w pierwszorzędowej korze wzrokowej [29].	20
2.12	Porównanie wyników otrzymywanych z modelu „integrate-and-fire” ze sprzężeniem i bez [29].	21
2.13	Pole rozszerzenia opisujące wpływ sąsiednich segmentów na jeden konkretny ze względu na (a) kąt oraz (b) siłę [10].	22
2.14	Wzajemny wpływ dwóch pól rozszerzenia na siebie i krzywa która jest na tej podstawie wydobywana [10].	23

2.15	Obrazy wejściowe dla modelu [10].	23
2.16	Otrzymywane wyniki przedstawione jako mapa istotności [10].	24
2.17	Wynik pracy modelu 2.17b z obrazem 2.17a przedstawiony w formie krawędzi. [10].	24
2.18	Model przedstawiający połączenia między elementami kory wzrokowej człowieka oraz jego podział na dwa szlaki: brzuszny i grzbietowy, które ściśle ze sobą są powiązane [25].	28
2.19	Struktura odzwierciedlająca hierarchię połączeń w modelu rozpoznawania obrazów [24].	29
2.20	Odzwierciedlenie zakresu szczegółowości na kolejnych poziomach modelu przedstawionego na rysunku 2.19 [25].	30
2.21	Przedstawienie szczegółów dotyczących podstawowych bloków, z których jest złożony model z rysunku 2.19 [26].	31
2.22	Kolejne kroki tworzenia modelu komponentów [17].	32
A.1	Nośnik, element funkcji Gabora, przedstawiony w przestrzeni i jako mapa.	35
A.2	Powłoka, element funkcji Gabora, przedstawiony w przestrzeni i jako mapa.	36
A.3	Funkcja Gabora przedstawiona w przestrzeni i jako mapa.	37
A.4	Wynik filtrowania obrazu A.5a pierwotną wersją filtru Gabora.	37
A.5	Wyniki filtrowania obrazu A.5a filtrami Gabora w których ω_0 jest w przedziale $(0, 1 >$	40
A.6	Wyniki filtrowania obrazu A.6a filtrami Gabora oraz operacji połączenia otrzymywanych rezultatów w jeden obraz.	41
A.7	Wyniki filtrowania obrazów A.7a i A.7d filtrami Gabora w których różna była wielkość okna filtru.	41

Spis tablic

Spis algorytmów

Bibliografia

- [1] Abbott L. F. Lapicque's introduction of the integrate-and-fire model neuron. *Brain Research Bulletin*, 50:303–304, 1999.
- [2] Csapo A., Roka A., Baranyi P. Visual Cortex Inspired Vertex and Corner Detection. *2006 IEEE International Conference on Mechatronics*, pages 551–556, 2006.
- [3] Curulku B., Lansner A. An Abstract Model Of A Cortical Hypercolumn. In *Proceedings of the 9th International Conference on Neural Information Processing (ICONIP'OP)*, volume 1, pages 80–85, 2002.
- [4] Downing P. E., Jiang Y., Shuman M., Kanwisher N. A cortical area selective for visual processing of the human body. *Science (New York, N.Y.)*, 293(5539):2470–3, 2001.
- [5] Garey L. Article reproduced by permission of the International Brain Research Organization. <http://www.korbinian-brodmann.de/english-article-mainmenu-12>, 1994.
- [6] Gauthier I., Behrmann M., Tarr M. J. Can face recognition really be dissociated from object recognition? *Journal of cognitive neuroscience*, 11(4):349–70, July 1999.
- [7] Girkin C. A., Miller N. R. MAJOR REVIEW Central Disorders of Vision in Humans. *Survey of Ophthalmology*, 45(5), 2001.
- [8] Grill-Spector K. The neural basis of object perception. *Current Opinion in Neurobiology*, 13(2):159–166, 2003.
- [9] Grill-Spector K., Kanwisher N. Visual Recognition, As soon as You know it is there, You know what it is. *American Psychological Society*, 16(2):152–160, 2005.
- [10] Guy G., Medioni G. Inferring Global Perceptual Contours from Local Features. *Contract*, (c):6–7, 1993.
- [11] Hubel D. H. Vision and the Brain At the January Stated Meeting David H . *Bulletin of the American Academy of Arts and Sciences*, 31(7):17–28, 4 1978.
- [12] Hubel D. H. Evolution of ideas on the Primary Visual Cortex, 1955-1978:A biased Historical Account. 12 1981.
- [13] Jones J. P., Palmer L. A. An evaluation of the two-dimensional Gabor filter model of

- simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1233–1258, 1987.
- [14] Kaas J. H. Progress and puzzles. *Current Biology*, 5(10):10–12, 1995.
 - [15] Kahana M. Virtual driving leads penn psychologists to the cells that sense direction in the brain: Path cells. 2010.
 - [16] Li Z. A neural model of contour integration in the primary visual cortex. *Neural computation*, 10(4):903–40, May 1998.
 - [17] Lopez P. M. Object component models using Gabor filters for visual recognition, September 2008.
 - [18] Movellan J. R. *Tutorial on Gabor Filters*, 2008.
 - [19] Ramachandran V., editor. *Encyclopedia Of The Human Brain*, volume 1. Elsevier Science, 2002.
 - [20] Resko B., Baranyi P. Lateral Operations in the VFA model. *IEEE*, pages 141–145, 2008.
 - [21] Resko B., Roka A., Baranyi P. Visual Cortex Inspired Intelligent Contouring. *IEEE*, pages 47–51, 2005.
 - [22] Riesenhuber M. *How a Part of the Brain Might or Might Not Work : A New Hierarchical Model of Object Recognition*. PhD thesis, Massachusetts Institute of Technology, May 2000.
 - [23] Schlei B. R. A new computational framework for 2D shape-enclosing contours. *Image and Vision Computing*, 27(6):637–647, 2009.
 - [24] Serre T., Kouh M., Cadieu C., Knoblich U., Kreiman G., Poggio T. *A Theory of Object Recognition: Computations and Circuits in the Feedforward Path of the Ventral Stream in Primate Visual Cortex*. 2005.
 - [25] Serre T., Poggio T., Riesenhuber M., Wolf L., Bileschi S. M. High-Performance Vision System Exploiting Key Features Of Visual Cortex. United states patent application publication, Massachusetts Institute of Technology, September 2006.
 - [26] Serre T., Wolf L., Bileschi S., Riesenhuber M., Poggio T. Robust object recognition with cortex-like mechanisms. *IEEE transactions on pattern analysis and machine intelligence*, 29(3):411–26, March 2007.
 - [27] Tadeusiewicz R., Korohoda P. *Komputerowa analiza i przetwarzanie obrazów*. Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków, 1 edition, 1997.
 - [28] Ungerleider L. G., Haxby J. V. 'What' and 'where' in the human brain. *Current Opinion in Neurobiology*, 4:157–165, 1994.
 - [29] Vanrullen R., Delorme A., Thorpe S. J. Feed-forward contour integration in primary visual cortex based on asynchronous spike propagation. *Neurocomputing*, 38:–401003–1009, 2001.
 - [30] Wiesel T. N. The Postnatal development of the Visual Cortex and the influence of environment. 12 1981.