
Aleksandar Sretenović i Vuk Vukomanović

Analiza karakteristika tekstura na fotografijama i njihova primena u prepoznavanju materijala

U ovom radu su ispitivane metode klasifikacija materijala fotografisanih pod nepoznatim uglom i osvetljenjem. Demonstrirano je kako se ova vrsta klasifikacije, gde je tekstura glavni vektor karakteristika, može uspešno izvršiti koristeći banke filtara. Konvolucijom slike sa određenom bankom filtara dobijeni su vektori karakteristika koji su potom klasterizovani, pa statistički obrađeni formiranjem histograma koji predstavlja raspored piksela po klasterima tog materijala. Histogrami su poređeni direktno, minimizacijom χ^2 distance, kao i korišćenjem neuronskih mreža, pri čemu je direktna metoda efikasnija (daje iste rezultate sa tri puta manjim trening setom). Poređene su efikasnosti različitih banki filtara. Ukupno je klasifikovano 20 materijala iz baze „Columbia-Utrecht” sa po 92 slike za svaki materijal.

Uvod

Klasifikacija je proces razvrstavanja podataka u klase, na osnovu zajedničkih osobina. U ovom radu klasifikacija se odnosi na razvrstavanje fotografija na osnovu materijala (teksture) koje one predstavljaju.

Pri klasifikaciji slika obično se izdvajaju karakteristike koje olakšavaju njihovo svrstavanje u jednu od klasa. Izdvajanje reprezentativnih karakteristika slike je često najveći izazov pri realizaciji kvalitetnog sistema za klasifikaciju. U fizičkom svetu, tekstura je okarakterisana bojom,

šarom, i reljefom, ali na slici takve karakteristike nije jednostavno izdvojiti. Dodatni je problem što se često dešava da ista tekstura slikana pod različitim uglovima i osvetljenjima izgleda sasvim drugačije, ili da dve sasvim različite teksture slikane pod izmenjenim uslovima izgledaju veoma slično.

Pokazalo se da se pogodne karakteristike fotografisane teksture mogu izdvojiti konvolucijom slike određenom grupom filtara (Varma i Zisserman 2004). Korišćeni filtri su predstavljani dvodimenzionalnim kvadratnim matricama. Osnovni princip primene ovih filtara je sledeći:

Neka su dati ulazna slika dimenzija $t \times t$ i filter dimenzija $r \times r$, pri čemu je svakom pikselu slike dodeljena samo jedna skalarna vrednost i gde je $t \gg r$. Postavimo centar filtra na proizvoljni piksel ulazne slike, tako da se svaki piksel filtra poklapa sa jednim pikselom slike. Množenjem vrednosti piksela filtra i odgovarajućeg piksela slike, pa kasnijim sumiranjem svih proizvoda, dobijamo vrednost koju nazivamo odziv filtra u konkretnom pikselu ulazne slike. Pomeranjem centra filtra kroz sliku i ponovnim računanjem dobijamo vrednosti odziva filtra za sve piksele ulazne slike. Novodobijena matrica je dimenzija $(t-r) \times (t-r)$ jer je neophodno da svaki piksel filtra poklapa neki piksel slike, pa centar filtra ne može biti bliže ivici slike od $r/2$ piksela.

Svaki piksel rezultujuće matrice nosi informaciju o odgovarajućem pikselu ulazne slike, kao i njegovoj $r \times r$ okolini. Tako se, uz pogodan izbor filtara, mogu posmatrati pojedinačni pikseli koji nose informaciju o lokalnim karakteristikama fotografisane teksture.

Aleksandar Sretenović (1994), Beograd, Miljkovačka 33g, učenik 4. razreda Matematičke Gimnazije u Beogradu

Vuk Vukomanović (1995), Kragujevac, 9. maja 61, učenik 3. razreda Prve kragujevačke gimnazije

MENTORI:

Đorđe Nijemčević, Microsoft Development Centar Serbia, Beograd

Miloš Stojanović, student Elektrotehničkog fakulteta Univerziteta u Beogradu

Ukoliko su filtri izabrani tako da ekstraktuju informaciju o teksturi na slici, sada imamo matricu $(t-r) \times (t-r)$ koja je nosilac te informacije. Međutim, matrica samostalno nije dovoljna za klasifikaciju, primarno zato što ne nosi dovoljan broj informacija za razlikovanje materijala.

Zato se prilikom filtriranja na sliku primenjuje skup (banka) filtara, koji imaju neke zajedničke osobine (npr. većina filtara je invarijantna na rotaciju, ili služi za detektovanje ivica). Nakon filtriranja slike bankom filtara, dobijamo n izlaznih matrica (odziva filtara) označenih sa M_1, \dots, M_n , gde je n broj filtara u banci. Označimo sa $M_i(p_j)$ vrednost i -tog odziva filtra za j -ti piksel ulazne slike, gde je $i = 1, \dots, n, j = 1, \dots, l$ i l ukupan broj piksela izlazne matrice. Na osnovu opisanih izlaznih matrica se formira l uređenih n -torki odziva svakog od filtara ($M_1(p_s), M_2(p_s), \dots, M_n(p_s)$) za svaki od $s = 1, \dots, l$ piksela izlazne matrice.

Sada se informacija o slici sadrži u skupu od l tačaka u n -dimezionom prostoru. Primenom metoda klasterovanja, konkretno K-means clustering (Moore 2001), te tačke se grupišu u k klastera (grupacija). Pikseli sličnih lokalnih svojstava teksture će, na ovaj način, biti grupisani u istom klasteru.

Dalje se svakom pikselu originalne slike dodeljuje vrednost iz opsega $1, \dots, k$, u zavisnosti od klastera kojem je taj piksel najbliži. Definišimo histogram H , koji prikazuje broj piksela dodeljenih svakom od klastera, kao $H(n_p(1), n_p(2), \dots, n_p(k))$, gde je $n_p(i)$ broj piksela dodeljenih klasteru $i, i = 1, \dots, k$. Formiranjem histograma H se dobija konačna informacija koja se koristi za određivanje teksture koja se nalazi na slici.

Metod

Realizacija se sastoji iz dva dela – treniranja i prepoznavanja materijala.

1. Treniranje

- Sve slike iz trening seta se filtriraju bankom filtara
- Nakon toga se za svaki materijal ponosob, i karakteristike izdvojene iz njega u vidu $(t-r) \times (t-r) \times n$ vektora, primenjuje K-means algoritam, kojim se dobija k

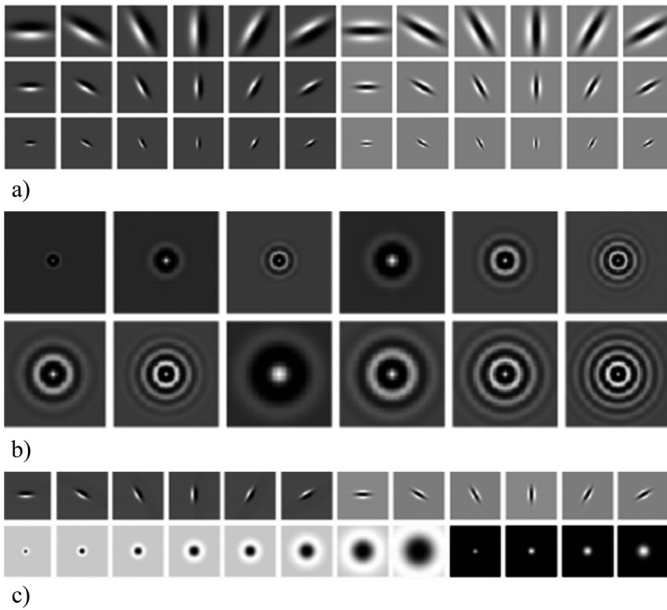
klastera i k njihovih centara odnosno srednjih vrednosti

- Pravljenje rečnika histograma za taj materijal, koji se potom koristi pri samom prepoznavanju
 - Ceo opisan proces se ponavlja za sve materijale
 - Na kraju treninga se dobija set od $L \times P$ histograma, gde je L broj klasa, a P broj slika u trening setu za jedan materijal
- ### 2. Prepoznavanje nepoznatog materijala
- Filtriranje neidentifikovane slike bankom filtara i računanje karakteristika u n -dimezionom prostoru
 - Za svaki od materijala nad kojima je izvršeno treniranje:
 - a) svrstavanje svakog od $(t-r) \times (t-r)$ vektora karakteristika uz jedan (najbliži) centar klastera, već izračunat za taj materijal u procesu treniranja
 - b) označavanje svakog piksela ulazne slike indeksom klastera kome njegov vektor karakteristika pripada
 - c) pravljenje histograma koji predstavlja raspored piksela po klasterima tog materijala
 - Poređenje histograma dobijenih iz nepoznate slike sa rečnikom histograma rezultuje identifikacijom materijala na nepoznatoj slici

Pre treniranja slike su normalizovane, kako bi klasifikacija u određenoj meri bila otporna na varijacije u kontrastu i osvetljenju.

Treniranje

Slike iz trening seta na kojima se nalazi isti materijal (pripadaju istoj klasi) filtriraju se izabranom bankom filtara. U ovom radu korišćene su tri banke filtara: MR8, LM i S (Varma i Zisserman 2004). MR8 banka filtara (slika 1a) se sastoji od orijentisanih filtara na različitim skalama, pri čemu se u obzir uzimaju samo filtri koji za datu orijentaciju daju maksimalni odziv. S banka (slika 1b) sadrži filtre slične high-pass filtrima (filtrar propusnik visokih učestanosti) u obradi signala, koji su rotaciono invarijantni (orijentacija šara na slici ne utiče na klasifikaciju). LM banka filtara (slika 1c) se sastoji od kombinacije orijentisanih filtara i simetričnih filtara.



Slika 1. Izgled korišćenih banki filtara: a) MR8 banka, b) S banka, c) LM banka.

Figure 1. Filter banks used in this paper: a) MR8 bank, b) S bank, c) LM bank.

Nakon filtriranja svih slika se dobija skup od $P \times (t-r) \times (t-r)$ tačaka u n -dimenzionom prostoru, pri čemu je P broj slika jedne klase u trening setu, t dimenzija slike (u ovom radu 200), a $r \times r$ dimenzije filtara. Svaka opisana tačka predstavlja jedan vektor karakteristika nekog od piksela ulazne slike. Dalje se na taj skup vektora karakteristika primenjuje K-means algoritam, čime se dobija k grupacija (klastera), odnosno k vektora koja predstavljaju centre ovih k grupacija.

Svakom opisanom vektoru karakteristika pridružuje se centar klastera kojem je on najbliži. Potom se za svaku sliku klase koja se obrađuje iz trening seta, tj. za skup tačaka u n -dimenzionom prostoru, koji se dobije posle filtriranja slike, odredi koliko tačaka pripada kom klasteru, tako što se kaže da tačka pripada klasteru ako je centar klastera bliži datoj tački od centara ostalih klastera. Time se dobije $(t-r) \times (t-r)$ matrica čiji su elementi brojevi između 1 i k . Poslednji korak je formiranje histograma H za svaku od slika. Histogram H prikazuje broj piksela dodeljenih svakom od klastera. Time se dobija P histograma koji predstavljaju deo rečnika koji opisuje koja slika pripada kojoj klasi.

Opisani proces se ponavlja za svaku klasu koju program treba da prepozna, i na kraju se za svaku klasu dobija po P histograma, kao i po k centara klastera. Važno je napomenuti da centri

klastera koji se dobijaju K-means algoritmom nisu uvek isti, i da umnogome zavise od rasporeda tačaka (nekada se ne dobijaju isti centri čak i kada se algoritam primeni dva puta na isti skup tačaka). Klasteri čiji se centri dobiju za svaku klasu se takođe nazivaju i tekstoni.

Na kraju treninga se dobija set od $L \times P$ histograma, gde je L broj klasa, i $L \times k$ tekstona.

Klasifikacija

Klasifikacija se vrši slično treningu, tako što se prvo slika koja se klasifikuje filtrira bankom filtara, čime se dobije skup tačaka u n -dimenzionom prostoru odziva filtara. Potom se, umesto primene K-means algoritma, za svaku klasu i njenih k centara klastera („naučenih” u procesu treninga) određuje koliko tačaka ulazne slike pripada kom klasteru, da bi se na kraju formirao histogram koji predstavlja raspodelu piksela slike po tekstonima date klase. Tako se dobija L histograma koji predstavljaju raspored piksela slike po tekstonima L klasa. Svaki od L histograma se upoređuje sa svakim od P histograma koji predstavljaju deo rečnika koji opisuje koja slika pripada kojoj klasi. Ona klasa među čijim se histogramima pronađe najveća podudarnost se histogramom nove slike je klasa kojoj pripada nova slika, tj. upoređivanjem histograma se određuje vrsta materijala na slici.

Analiza efikasnosti poređenja histograma

Jedan od ciljeva ovog rada je poređenje efikasnosti metoda upoređivanja histograma. Konkretno, za poređenje histograma korišćene su metoda direktnog upoređivanja minimizacijom χ^2 distance i primenom klasifikacije putem neuronskih mreža. Takođe, broj klastera na koji se grupišu tačke utiče na efikasnost klasifikacije (premalu klastera uzrokuje nedovoljnu definisanost klasa, dok se sa previše klastera može desiti da i najsitnije razlike svrstaju dve slične slike u različite klase). U nastavku će biti opisane korišćene metode upoređivanja histograma.

Metod direktnog poređenja histograma koristi minimizaciju χ^2 distance. Klasa u kojoj se nalazi histogram za koji ova vrednost hi-kvadrata najmanja, je tražena klasa.

Metod treniranja i korišćenja neuronskih mreža podrazumeva treniranje onoliko neuronskih mreža koliko imamo klasa koje sistem prepoznaje. Svaka mreža se trenira da samo za jednu klasu prepozna da li slika pripada toj klasi ili ne. Ulazni parametri su u ovom slučaju histogrami, a izlazni parametar je vrednost, označena sa P_r , koja predstavlja verovatnoću da slika pripada klasi čija se neuronska mreža testira. Prilikom klasifikovanja korišćenjem opisane metode, L različitih histograma nove slike (po jedan histogram za svaku klasu) propušta se kroz L različitih neuronskih mreža (jedna mreža za svaku klasu), pri čemu i histogram i mreža odgovaraju istoj klasi. Na kraju se poredе dobijene vrednosti verovatnoća P_r i klasa za koju je verovatnoća P_r najveća se proglašava klasom kojoj nova slika pripada.

Rezultati i diskusija

Uspešnost metode direktnog poređenja histograma u zavisnosti od broja klastera i broja slika u trening setu data je u tabelama 1, 2 i 3, za S, MR8 i LM banke filtara, respektivno. U svim slučajevima klasifikacija se vršila na 20 različitih klasa.

Iz priloženih tabela se može zaključiti da je i veličina trening seta, kao i broj klastera u mnogome utiču na efikasnost sistema. Ostaje još da se ispita koliko broj klasa otežava klasifikaciju.

Tabela 1. Uspešnost klasifikacije 20 različitih klasa u zavisnosti od broja klastera i veličine trening seta, koristeći S banku filtara

Broj slika u trening setu	Broj klastera		
	10	20	25
20	81%	88%	89%
15	72%	79%	/

Tabela 2. Uspešnost klasifikacije 20 različitih klasa u zavisnosti od broja klastera i veličine trening seta, koristeći MR8 banku filtara

Broj slika u trening setu	Broj klastera		
	10	20	25
20	83%	92%	/
15	75%	90%	/

Tabela 3. Uspešnost klasifikacije 20 različitih klasa u zavisnosti od broja klastera i veličine trening seta, koristeći LM banku filtara

Broj slika u trening setu	Broj klastera		
	10	20	25
20	80%	86%	86%
15	78%	83%	/

Prilikom rada sa neuronskim mrežama, vršena je klasifikacija na 20 različitih klasa, tj. trenirano je 20 različitih neuronskih mreža, za svaki materijal po jedna. Svaka mreža je imala 25 ulaznih neurona (po jedan za svaki klaster predstavljen na histogramu), 25 skrivenih, i jedan izlazni. Takođe, važno je napomenuti da su mreži prosleđene sve slike koje se nalaze u bazi, od kojih je odabrano 70% za trening, a po 15% za verifikaciju i testiranje. Tokom provere efikasnosti, mreži su prosleđivane sve slike iz baze, koje pripadaju jednoj od 20 klasa. Tačnost opisanog sistema iznosila je 79%. Iz navedenog se

Tabela 4. Uspešnost klasifikacije 20 klasa u zavisnosti od broja klastera i veličine trening seta, koristeći S banku i neuronske mreže

Broj slika u trening setu	10 klastera 10 ulaznih neurona	20 klastera 20 ulaznih neurona	25 klastera
20	22%	53%	/
15	13%	29%	/

može uočiti da je neuralnoj mreži za istu tačnost potreban mnogo veći trening set nego kada se vrši direktno poređenje. U tabeli 4 prikazana je uspešnost neuronskih mreža pri klasifikaciji koristeći S banku filtara i redukovan trening set.

Zaključak

Ovim radom smo demonstrirali kako je moguće izvršiti klasifikaciju slika fotografisanih pod nepoznatim uglom i osvetljenjem. Pokazano je kako se korišćenjem banki filtara može uspešno ekstrahovati informacija o teksturi iz slike. Istražena je zavisnost broja klastera, banki filtara i broja slika u trening setu od uspešnosti sistema. Takođe su upoređena dva načina prepoznavanja slika, tj. poređenja histograma – korišćenjem minimizacije χ^2 distance (direktno) i treniranjem neuralnih mreža. Pokazalo se da je za istu uspešnost neuronskim mrežama potreban tri puta veći trening set.

U daljim radovima zanimljivo će biti istražiti kako bi se karakteristika boje mogla iskoristiti u cilju poboljšanja sistema.

Literatura

Davies E. R. 2012. *Computer and Machine Vision: Theory, Algorithms, Practicalities*. Academic Press

Fukunaga K. 1990. *Introduction to Statistical Pattern Recognition*. Academic Press

Moore A. 2001. K-means and Hierarchical Clustering – Tutorial Slides. Department of Computer Science, Carnegie Mellon University. Dostupno na: <http://www.autonlab.org/tutorials/kmeans11.pdf>

Pele O., Werman M. 2010. The Quadratic-Chi Histogram Distance Family. U *Proceedings of The 11th European Conference on Computer Vision, ECCV 2010, Heraklion, Hellas* (ur. K. Daniilidis, P. Maragos i N. Paragios), Part II. Springer, str. 749-762.

Varma M., Zisserman A. 2005. A Statistical Approach to Texture Classification from single image. *International Journal of Computer Vision*, **62** (1/2): 61.

Zisserman A., Varma M. 2002. Classifying Images of Materials: Achieving Viewpoint and Illumination Independence. U *Proceeding of the 7th European Conference on Computer Vision, ECCV 2002, Copenhagen, Denmark*, (ed. A. Heyden, G. Sparr, M. Nielsen i P. Johansen), Part III. Springer, str. 255-271.

Aleksandar Sretenović and Vuk Vukomanović

Feature Analysis of Image Texture and their Application on Classification of Materials

The subject of this paper is material classification from images obtained under unknown viewpoint and illumination. In practice, it has been shown that this type of image classification, where the main “feature” vector is texture, can be successfully done using banks of filters. A filter bank is actually a set of filters with a common purpose (for example, the S-bank of filters contains filters whose purpose is detecting edges). The image is first transformed, in such a way that all of the pixels contain only one value, usually into the grayscale format. Afterwards, all of the

images of the same material are convolved with the chosen filter bank, leaving us with a matrix in which every pixel carries information about itself and its surroundings. Those matrices are clustered, using K-means clustering algorithm, into n clusters, which represent the features of a texture. Doing this for all images of a certain texture, and afterwards creating a histogram of clusters for each image, gives us the characteristics of that texture. While classifying a certain image, using the above explained method, a histogram is created for that particular image, and is afterwards compared to all the other histograms of all the textures. The texture whose histogram is the closest match to the histogram of the image should be the texture that is on the image.

In this paper, the efficiency rates of different banks of filters (the S-bank, MRS-bank and the LM-bank) of filters are compared, as well as the efficiency rates of two different methods of comparing histograms (the minimisation of chi-squared distance and the usage of neural networks) are investigated. It is also investigated how much the number of clusters during the K-means clustering affects the effectiveness of the system (Moore 2001).

The results, when classifying 20 different materials, show that the number of images in the

training set greatly influences the efficiency rate, while the number of clusters influences it, but after a certain point it starts being less and less important. As shown in Table 1, the results of the system are almost 10% better with 20 training images (81.5% with 10 clusters, and 88.5% with 20 clusters) than with 15 (72% with 10 clusters, and 79.5% with 20 clusters). At the same time, while keeping the number of training images at 20, the system has performed 7% better with 20 clusters (88.5%) than with 10 (81.5%), but only 0.5% better with 25 clusters (89%), compared to 20. In all of these cases, the minimisation of chi-squared distance for comparing histograms and the S-bank of filters are being used (Pele & Werman 2010).

Surprisingly enough, while comparing the efficiency of neural networks and the chi-squared method, it is noticed that neural networks need a training set almost three times larger than the chi method, for a remotely close, but still lower efficiency rate. While the efficiency of the chi-squared method with the S-bank of filters and 25 clusters when classifying 20 textures is 89%, the efficiency of neural networks in the same system, but with three times more images in the training set is only 79.5%. 