

Projekt: Identifiering av handskrivna siffror

Kravspecifikation

Fredrik Berntsson

Version 1.0

Status

Granskad	FB	2022-06-10
Godkänd	FB	2020-06-10

Dokumenthistorik

Version	Datum	Utförda ändringar	Utförda av	Granskad
1.0	2022-06-10	Första versionen	FB	FB

1 Introduktion

I flera sammanhang finns det behov av att automatiskt läsa text. Ett exempel är fakturor där kontonummer kan läsas maskinellt. De flesta algoritmer för automatisk läsning av text kräver att det tecken som skall läsas finns tillgängligt som en bild i digital form. I projektet skall vi studera metoder för att automatiskt identifiera tecken som lagrats som bilder.

1.1 Parter

Med leverantör åsyftas i efterföljande text projektgruppen. Matematiska Institutionen (MAI) är beställare och kund.

1.2 Syfte och Mål

Målet med projektarbetet är att inhämta kunskap i hur matematiska metoder kan användas i realistiska tillämpningar, samt inhämta tillräcklig kunskap inom tillämpningen bildbehandling för att förstå vilka möjligheter och begränsningar som metoderna som studeras har. Projektet har också som övergripande mål att ge träning i att utveckla ett system, ingenjörsmässiga arbetsmetoder, samt dokumentation och muntlig presentation.

1.3 Användning

Beställaren skall kunna använda systemet för att lösa ett antal olika uppgifter. Systemet skall utformas så att det är enkelt för användaren att vidareutveckla.

2 Systemöversikt

Kärnan i systemet skall vara ett antal funktioner som tillsammans löser identifieringsproblemet. De bilder som används antas vara 16×16 pixlar. Ett antal olika metoder skall implementeras och testas i projektet.

Systemet skall använda en *referensmängd* med siffror av känd typ för att avgöra vilken typ av siffra en okänd bild föreställer. De siffror som skall användas finns lagrade i en Matlab fil `DataSet.mat`. De enskilda siffrorna finns lagrade i matriser där varje kolumn representerar en specifik siffra. Referensmängden finns lagrad i matrisen `RefSet` och en det finns även testmängd `TestSet` som man kan använda för att utvärdera metoderna. Det finns även två vektorer `RefAns` och `TestAns` som innehåller information om vad referensmängden respektive testmängden innehåller.

2.1 Avgränsningar

Systemet skall ses som en prototyp och inget högre krav ställs på användargränssnitt. Det räcker om systemet består av ett antal fristående funktioner som tillsammans kan användas för att lösa uppgiften. Det måste isåfall medfölja ett tydligt exempel där det visas hur de ingående funktionerna skall användas tillsammans för att lösa bildbehandlingsproblemen.

2.2 Generella krav på hela systemet

Krav 1	Original	Systemet skall kunna typbestämma okända siffror.	Bas
Krav 2	Original	Det skall finnas möjlighet att visa siffror grafiskt på skärmen.	Bas
Krav 3	Original	Systemet skall vara testat så att det är så stabilt som möjligt.	Bas

3 Metoder för teckenidentifiering

I projektet representeras varje siffra av en kolumnvektor. Vi kan därför mäta skillnaden mellan två siffror genom att beräkna det Euklidiska avståndet mellan dem. Närmsta-granne-metoden utnyttjar detta för att typbestämma siffror. Givet en okänd siffra och en referensmängd med kända siffror beräknar vi avståndet mellan den okända siffran och siffrorna ur referensmängden. Vi hoppas sedan att den okända siffran är av samma typ som den siffra ur referensmängden som ligger närmast.

Krav 4	Original	Det euklidiska avståndet mellan två vektorer, som representerar siffror, skall kunna beräknas.	Bas(1)
Krav 5	Original	Systemet skall kunna hitta den siffra ur referensmängden som ligger närmast en viss okänd siffra.	Bas
Krav 6	Original	Beskriv matematiskt hur avståndet mellan två siffror skall beräknas.	Bas(1)
Krav 6	Original	Beskriv matematiskt hur avståndet mellan två siffror skall beräknas.	Bas(1)
Krav 7	Original	Beskriv schematiskt hur Närmsta-granne metoden skall implementeras.	Bas(1)

Det finns fler avståndsmått att välja på än just Euklidiskt avstånd. En mer abstrakt matematisk definition av begreppet avstånd kan hittas i läroböcker. Ett exempel på ett mer generellt avstånd är p -normen där skillnaden mellan två vektorer beräknas som

$$\|x - y\|_p = \left(\sum_{k=1}^n |x_k - y_k|^p \right)^{1/p},$$

där olika värden på p ger olika avståndsmått. En naturlig fråga blir då om andra avståndsmått kan ge bättre resultat än det Euklidiska avståndet.

Krav 8	Original	Hitta en referens där en matematisk definition av avstånd mellan två vektorer ges. Visa sedan att såväl det Euklidiska avståndet, samt p -normen, för vissa värden på p , är avståndsmått.	Bas(1)
Krav 9	Original	Implementera Närmsta-granne-metoden så att p -normen kan användas vid avståndsberäkningen. Jämför även träffsäkerheten för några värden på p med det Euklidiska avståndet (alltså fallet $p = 2$).	Bas
Krav 10	Original	Visa att fallet $p = 1/2$ inte ger ett avståndsmått enligt definitionen. Använd trots detta $p = 1/2$ i avståndsberäkningen och se hur bra resultat som fås. Pröva även några fler otillåtna värden på p .	Extra

Referensmängden innehåller ett stort antal siffror av varje typ. Ett effektivt sätt att utnyttja siffrorna i referensmängden kan utvecklas om man betraktar siffrorna av en viss typ som vektorer i ett linjärt rum. Vi kan då beräkna en ortogonalbas för detta rum på följande sätt: Bilda den första basvektorn genom att bilda medelvärdet av alla siffror av en viss typ. Använd sedan Gram-Schmidt ortogonalisering för att göra alla vektorer ortogonala mot denna första basvektor. Första basvektorn representerar då det mest typiska utseendet för en siffra av denna typ. Andra basvektorn fås sedan genom ytterligare en medelvärdesbildning över vektorerna. Den andra basvektorn kommer då att representera den mest typiska variationen i hur den aktuella siffertypen skrivs. Man kan på samma sätt fortsätta generera fler basvektorer genom att fortsätta Gram-Schmidt ortogonalisering och medelvärdesbildning.

Istället för att beräkna avståndet mellan en okänd siffra och enskilda siffror ur referensmängden kan vi beräkna avståndet mellan en okänd siffra och det underrum som spänns upp av k st basvektorer skapade enligt metoden ovan. Avståndet kan definieras genom att den okända siffran projiceras på underrummet. Denna metod kan bli mycket effektivare än närmsta-granne-metoden förutsatt att det räcker med ett litet antal basvektorer för att fånga de vanligaste sätten siffror kan skrivas på.

Notera att basvektorerna inte behöver beskriva den okända siffran särskilt väl. Det viktiga är att de basvektorer som skapats från rätt siffertyp är bättre än övriga uppsättningar basvektorer.

Krav 8	Original	Beskriv de matematiska formler som skall utnyttjas då en ortogonalbas skall beräknas enligt metoden som beskrivits ovan.	Bas(1)
Krav 9	Original	Beskriv matematiskt de formler som behövs för att beräkna avståndet mellan en vektor (okänd siffra) och det underrum som spänns upp av k st ortogonala basvektorer.	Bas(1)
Krav 10	Original	Givet en uppsättning vektorer skall systemet kunna beräkna en ortogonal bas enligt metoden ovan. Användaren skall kunna bestämma hur många basvektorer som efterfrågas.	Bas
Krav 11	Original	Systemet skall kunna beräkna avståndet mellan en vektor, som representerar en okänd siffra, och det linjära underrum som spänns upp av ett antal basvektorer.	Bas
Krav 12	Original	Systemet skall med hjälp av tidigare beräknade linjära underrum avgöra av vilken typ en okänd siffra är.	Bas
Krav 13	Original	Systemet skall kunna beräkna projektionen av en testsiffra på ett underrum samt visa denna grafiskt.	Extra

4 Tillämpning och Utvärdering

För att utvärdera hur de metoder som utvecklats fungerar finns en testmängd. Genom att försöka typbestämma siffrorna ur testmängden och jämföra med det facit som finns så kan vi se ifall metoderna fungerar bra.

Krav 13	Original	Försök att typbestämma alla siffror i testmängden med hjälp av Närmsta-granne-metoden. Räkna hur stor andel som klassas korrekt.	Bas
Krav 14	Original	Försök att typbestämma alla siffror i testmängden med hjälp av underrums-metoden. Använd $k = 15$ basvektorer av varje siffertyp. Ränka efter hur stor andel som typbestämms korrekt.	Extra
Krav 15	Original	Låt parametern k variera i underrums-metoden och se efter hur metodens träffsäkerhet beror på antalet basvektorer som används.	Extra
Krav 16	Original	De olika siffertyperna kan vara olika svåra att typbestämma korrekt. Undersök hur stor andel fel som fås för varje siffertyp.	Extra

5 Ekonomi

Krav 17	Original	Projektet skall genomföras med en arbetsinsats på 105 timmar per student.	Bas
----------------	-----------------	---	------------

6 Leveranser

Vid slutleverans skall samtliga krav märkta **Bas** vara uppfyllda. För grupper med 5 studeneter gäller att minst ett krav märkt **Extra** dessutom skall vara uppfyllt. För grupper med 6 studenter gäller att tre **Extra** krav skall vara genomförda. För större grupper gäller att samtliga **Extra** krav skall vara uppfyllda.

Krav 18	Original	Leverans av gruppkontrakt skall ske till handledare och kursansvarig.	Bas
Krav 19	Original	Leverans av projektplan skall ske via e-post till beställaren och handledaren.	Bas
Krav 20	Original	Presentation av krav Bas(1) skall ske muntligt för beställare. Detta sammanfaller med BP3.	Bas
Krav 21	Original	Delleverans omfattande det utvecklade systemet, bruksanvisning, och ett utkast till den konferensartikel som presenterar arbetet skall skickas till beställare och handledare.	Bas
Krav 22	Original	Slutleverans bestående av den artikel som beskriver gruppens arbete skall skickas till handledare och beställare. Det skall tydligt framgå i texten att alla Bas-krav skall är uppfyllda samt extrakrav enligt texten ovan.	Bas
Krav 23	Original	Statusrapporter skall skickas via e-post till beställare varannan måndag under tiden som projektarbetet pågår.	Bas
Krav 24	Original	Tidsrapporter för varje vecka skall skickas via e-post till handledare och beställare senast måndag 13.00 påföljande vecka	Bas
Krav 25	Original	Leverans av efterstudie skall ske till kursansvarig.	Bas
Krav 26	Original	Krav märkta Bas* skall vara uppfyllda vid kursens avslutande och redovisning av detta sker genom mail till beställaren.	Bas

7 Dokumentation

För att beställaren skall kunna testa det utvecklade systemet behövs en bruksanvisning. Denna bruksanvisning skall förklara hur man gör för att använda systemets olika funktioner.

Krav 27	Original	En bruksanvisning skall följa med systemet.	Bas
Krav 28	Original	En projektplan skall upprättas	Bas
Krav 29	Original	Mötesprotokoll skall föras vid alla möten.	Bas*
Krav 30	Original	Tidsrapporter skall sammanställas varje vecka.	Bas*
Krav 31	Original	Projektgruppensarbete skall beskrivas i en artikel som lämpar sig för att presenteras på en teknisk konferens. Artikeln skall innehålla nödvändig bakgrundsinformation, en beskrivning av vad som gjorts, samt exempel som illustrerar hur metoden fungerar. Detta dokument skall inte följa LIPS mallarna.	Bas*
Krav 32	Original	En litteraturstudie skall göras och några artiklar som beskriver liknande, eller alternativa, tillvägagångssätt skall identifieras. Detta för att beställaren skall kunna utveckla projektet vidare.	Bas*
Krav 33	Original	En efterstudie skall skrivas.	Bas*
Krav 34	Original	Det skall skrivas en text om hur det utvecklade systemet kan bidra till ett hållbart samhälle. Omfattningen bör vara en till två paragrafer. Denna skall ingå i artiklens introduktion eller i sammanfattningen.	Bas*
Krav 35	Original	Alla dokument som lämnas till beställare skall granskas, med avseende på både språk och innehåll, av minst en projektdeltagare, utöver den som skrev texten.	Bas*
Krav 36	Original	Alla dokument skall följa LIPS-mallarna	Bas*