

Knut Kleve og Espen Smith Ore
Universitetet i Oslo

"PATTERN RECOGNITION" I PAPYRUSFORSKNING

Problemet er å rekonstruere bokstaver som det bare finnes blekkrester igjen av på papyrusunderlaget. Rekonstruksjon ved hjelp av fotografiske metoder har vist seg nytteløs. Der hvor det ikke er blekk, er skriften totalt forsvunnet, uten å etterlate seg spor. Rekonstruksjon ved hjelp av EDB-metoder gjenstår.

Et bokstavfragment kan sammenlignes med modeller av fullstendige bokstaver (jfr. fig. 1). Hvis de forskjellige bokstaver alltid hadde vært likt skrevet, ville saken være enkel. Opplysninger om de fullstendige modeller kunne lagres på en slik måte at de og fragmentene uten videre kunne sammenlignes på en datamaskin. Bokstaver er figurer som kan beskrives i en så finmasket todimensjonal matrise som det er ønskelig eller praktisk mulig å bruke.

Bokstavene i papyrene er imidlertid håndskrevne og varierer i størrelse, sentrering i bildet, helningsvinkel osv. Derfor er det ikke nok å bruke et program som bare ser etter om en figur (fragmentet) kan være en del av én av en gitt mengde større figurer (modellbokstavene). Man må statistisk avgjøre hva som er de mest typiske kjennetegn ved bokstavene i en papyrustekst og prøve å finne tilsvarende kjennetegn i det aktuelle bokstavfragmentet.

Slike statistiske gjennomsnittsbokstaver kan brukes til å sammenligne bokstavene som figurer. Det må da legges inn et slingringsmonn når det skal avgjøres om figurer er like (jfr. sammenligning av "fuzzy sets").

Det er mest sannsynlig at den endelige avgjørelse etter sammenligningen av fragment og modeller må tas av et menneske. Tegnene kan lett variere så meget at det ikke vil være mulig for maskinen å avgjøre hvilken bokstav fragmentet er en del av (jfr. fig 2). men man kan tenke seg et program som svarer med et utvalg av mulige bokstaver etter at fragmentet er sammenlignet med listen av modeller. Deretter kan man selv sammenligne fragmentet med de alternativer maskinen har gitt.

Sammenligningen kan foregå på følgende måte: Ved hjelp av en grafisk terminal legger man bilder av fragmentet og de forskjellige modellene over hverandre, og benytter even-

tuelte også et program som kan forstørre, forminske og vri bildene.

Input/output

Innlesningen av bokstavene og fragmentene kan foregå på forskjellige måter:

Man kan selv tegne et rutenett over et (forstørret) fotografi av tegnet.

Man kan la maskinen lese inn tegnet direkte fra et fotografi ved hjelp av optisk leseutstyr.

Man kan bruke et digitaliseringsbord hvor man tegner over et bilde av tegnet i ønsket størrelse med en magnet-penn som sørger for at bildet blir lagret med den nøyaktighet vi ønsker.

Utmatingen av (rekonstruerte) bokstaver kan skje grovt med en tegnmatriks på en linjeskriver, eller man kan få det tegnet på en plotter.

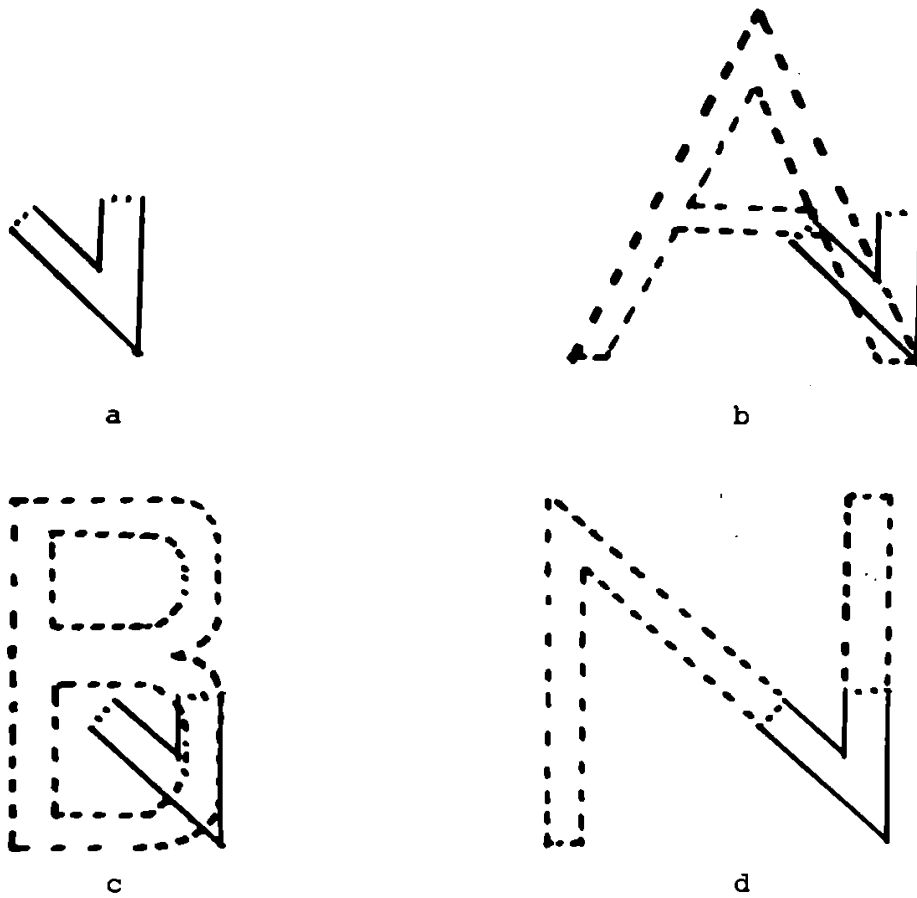


Fig. 1. a: fragmentet, b og c: sammenligning med modeller som ikke passer, d: fragmentet er en del av denne modellen.



Fig. 2. Her er det større likhet mellom modellbokstaven (prikket linje) A og bokstaven R enn mellom modell-A og skrevet A.