Location of structural sections from within a highly distorted complex line drawing

R.O. Canham, S.L. Smith and A.M. Tyrrell

Abstract: Complex and highly distorted line drawings are produced by subjects attempting the Rey Complex Figure test, a clinical test of neuropsychological assessment. However, the marking scheme conventionally employed can be subjective and unreliable. In this paper, the first stages in automating this scoring system are investigated using a robust technique to locate a reduced set of scoring sections and a knowledge-based system that employs spatial metrics and fuzzy approximation techniques. Testing the technique using clinical data produced encouraging results that support the argument that this is a feasible approach for implementing a fully automated system, and that in its current state, can be immediately applied in a semi-automated system.

1 Introduction

In this paper we report progress towards automated scoring of the Rey Complex Figure (RCF) test [1, 2], a clinical test of neuropsychological assessment, for which subject responses are typically complex and highly distorted line drawings. Through robust estimation of the distorted figural sections and spatial aggregation by fuzzy approximation, our methods solve significant problems in interpretation that clinicians currently have to do by hand. We demonstrate how these methods might be applied in a fully-automated system, identifying modules required to complete such a system and the research challenges they present, but we also show how our methods can be immediately applied in a semi-automated context to reduce clinician workload.

The RCF test is used to evaluate visual perception and long term visual memory. It has been chosen for automation for two reasons: first, it is probably the most complex neuropsychological test of its type and as such provides a suitable challenge by which the success of automating tests of this type can be measured, and secondly, the RCF is widely used in neurological assessment and consequently, it is hoped this work will be of interest and benefit to the medical community.

A typical protocol is to present the figure (as shown in Fig. 1) to the subject, who attempts to reproduce it as accurately as possible. The figure and the subject's attempt are removed and the subject is then asked to produce a copy from memory.

A commonly used scoring system, devised by Osterrieth, considers 18 sections (shown in Fig. 2) which are assessed by their location and degree of distortion. Currently this process is performed by hand, in a subjective manner, which is open to interpretation [3–5].

Automation of the scoring system would provide an objective and consistent result, while alleviating a highly skilled clinician from a tedious and time consuming task.

In this paper we consider the location of scoring sections within the subject's drawing, from those shown in Fig. 2. This is the second and most demanding stage in the full automation of the Rey Complex Figure test. The first stage, the identification of simple geometric shapes as candidates for these scoring sections, has previously been demonstrated by the authors [6]. The third and final stage is the assignment of a numerical score for these scoring sections, which, by comparison to the previous stages, is relatively straightforward.

The nature of the Rey Complex Figure test is such that the data it generates can be extremely distorted. Scoring sections can be misplaced, repeated, or missing; they can be incorrectly sized with large gaps in lines at corners or along a side; lines can be bent, twisted, stepped, curved and deviate a long way from a straight path. The sections can be wildly out of proportion and lines can be multi-stroke. Figure 3 shows distortions present in clinical data used for testing the system.

Identification of a particular section is achieved through the identification and location of its neighbouring sections. However, it is difficult to determine whether it is the section being considered that is incorrectly positioned, or the section it is being compared with.

As mentioned above, we have previously demonstrated that it is possible to identify the fundamental shapes that are candidates for one of the scoring sections (specifically triangles, rectangles and diamonds) using a perceptual grading employing fuzzy techniques and Gestalt psychology [6]. The scoring sections can then be located using relative spatial metrics between the candidates. This is a computationally expensive process and so we reduce the number of candidates using absolute spatial metrics [7].

2 Previous work

We have failed to identify any previous attempts to automate the RCF scoring system in the literature, and

IEE Proceedings online no. 20045166

doi: 10.1049/ip-vis:20045166

Paper first received 7th September 2004 and in revised form 9th March 2005

The authors are with Department of Electronics, University of York, Heslington, York, YO10 5DD, UK

E-mail: sls5@ohm.york.ac.uk

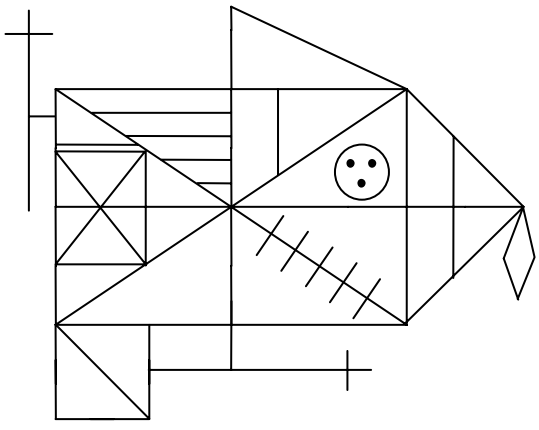


Fig. 1 *The Rey complex figure*

found few examples detailing the automation of other similar pen and paper tests. The latter is largely confined to figure copying of simple geometric shapes such as cubes [8, 9], squares and crosses [10, 11] for diagnosis of a range of neurological conditions including stroke, Alzheimer's disease and Parkinson's disease.

Hand drawn line figures and sketches are generated in a number of applications and appeared to have an obvious similarity to the RCF. The more complex of these recognise hand drawn symbols for circuit and flow diagrams [12, 13] and hand drawings [14, 15]. However, the symbols are typically identified by their constituent parts, rather than the juxtaposition and relationships with their neighbours and thus, are not suitable for the RCF owing to the possibility of missing and incorrectly positioned sections.

The conversion of paper plans and drawings for use within CAD systems produces a complex line drawing that can be damaged, introducing distortion and hence, has a number of similarities to the RCF. Examples include [16, 17], which typically use a knowledge based system [18]. Graph matching [19] and relaxation probabilities [17] have also been employed but again simple absolute thresholds and consistency of rules are employed respectively.

Many computer vision and image understanding techniques use edge detection and other pre-processing techniques to reduce a complex image into a number of significant lines. Distortion and occlusion are often present producing a scenario which is actually closer in nature than the processing of hand drawn figures and conversion of paper plans described above. Applications include robot navigation [20, 21], building location [22, 23] and scene interpretation [24, 25]. The spatial relationships used to locate features are generally intuitive and heuristic in nature to describe relationships such as near, above, below, left and right. With a common junction, relationships such as above and below can be very simply defined, however, with more complex areas these relationships are no longer simple. Keller and Wang [26], and Krishnapuram and Ma [27] compared three spatial relation techniques; the aggregation, centroid and compatibility methods. Experimental results indicated that the aggregation method gave more intuitive results.

The human visual system can easily group certain elements together that are considered as visually important. Gestalt Psychology [28, Chapter 5] proposes that the properties of the whole are not the result of the summation of the parts and its relevance to visual perception is considered in this work. The features identified as significant in the grouping procedure can be summarised as proximity, curvilinear continuation, similarity, closure, symmetry, common region and connectedness [29].

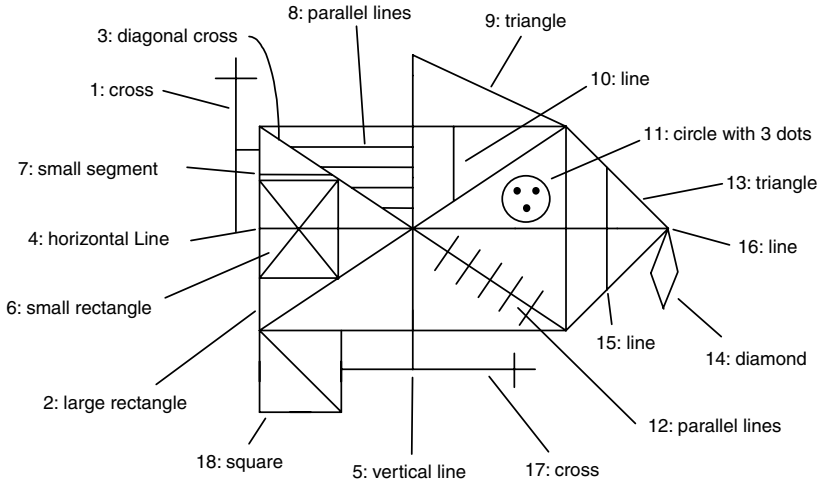


Fig. 2 *The Osterrieth scoring system*

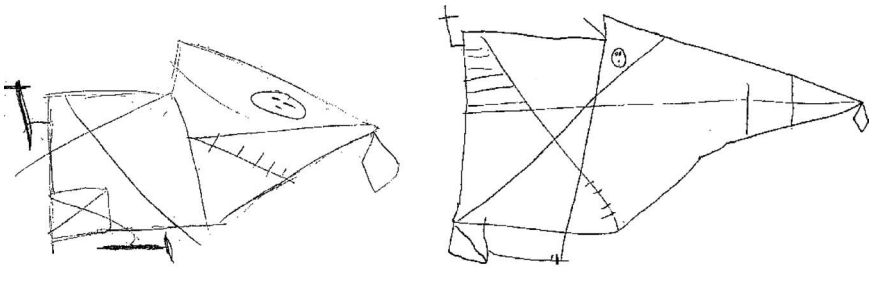


Fig. 3 *Examples of distorted figures*

The range of applications that have similarities to the RCF is large but none have the same degree of distortion. A flexible and robust technique is required to describe the perceptual significance of the features to be identified.

3 Description of system

We scan a subject's test responses and after pre-processing vectorise constituent lines to form a frame system. This is comprehensively searched to find the base shapes using a rule set for a corner and continuation of a straight side. Located shapes are then rated on perceptual distortion using fuzzy metrics.

The number of candidates for each scoring section is reduced by initially considering each candidate in isolation using unary metrics. These metrics are fuzzy in nature and based upon absolute position, size, orientation and basic shape features. Each candidate for each section has a similarity metric calculated based upon the aggregate of a number of these metrics. Features with a measure above a working threshold are considered for further processing.

With the number of candidates for a section reduced we apply fuzzy binary metrics based upon Gestalt features such as relative position, size, contiguous, touching, inside, outside, above, below, left, right and relative orientation. Each section has a number of metrics defined, which are calculated and aggregated to give a final measure.

Since the RCF has eighteen scoring sections we defined a number of smaller subsections of the figure to prevent combinatorial explosion. The best candidates for the subsections are combined to generate the scoring section instances for the whole figure.

We now describe the system implemented to identify relevant scoring sections within the RCF in more detail by considering the following sub-processes: candidate reduction, binary relations and figure aggregation.

3.1 Candidate reduction

We reduce the possible scoring section candidates from the features found by use of unary and non-specific binary metrics. The unary metrics are fuzzy in nature and based upon absolute position, size, orientation and basic shape features. The non-specific binary metrics are largest size, extreme point and distinct feature. The largest size metric compares a feature's size to the largest of all the candidates for that section. This is expressed as a fuzzy membership by using the half trapezoidal function, μ_h [30]. If the set of all the candidates for a given scoring section is *SCORE* and the size of a candidate, $c_i \in SCORE$, is denoted by s_i then the largest size metric is defined by (1).

$$\mu_{largest}(c_i) = \mu_h\left(\sqrt{\frac{\max_{c_j \in SCORE[s_j]}}{s_i}}\right) \quad (1)$$

The extreme point metric compares, in a given axis, all the points of a candidate shape with the extreme points of all the candidates for that section. This metric reduces the effects of other misplaced sections, as shown in Fig. 4. Although the cross is much higher than triangular Section 13, this triangle still has the upper point above all other candidates, which is the desired result. This is described in (2)

$$\mu_{extreme}(c_i) = \mu_h\left(\frac{\min_{p_j \in P_i} \min[|p_j - ep1|, |p_j - ep2|]}{s_i}\right) \quad (2)$$

where the set of all corner positions of a candidate in a given plan is P_i and $ep1$ and $ep2$ are the two extreme corner locations of all the candidate shapes. s_i is the size of the figure in the appropriate axis.

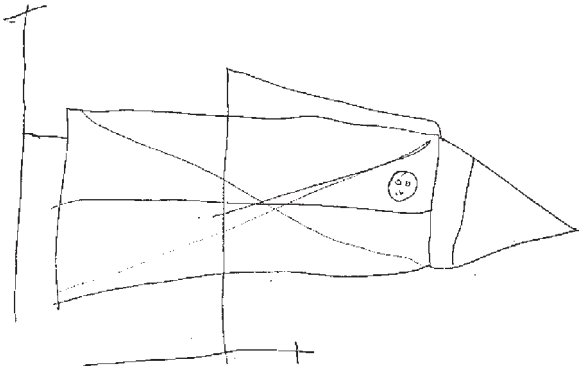


Fig. 4 *Example of effectiveness of the extreme point metric*

The highlighted triangle in Fig. 5 can best be described as part of the rectangle which has a diagonal line through it. The scoring sections are currently considered in isolation and so this would be considered as a perfect triangle. The distinct metric grades triangles that are within rectangles. A triangle is indistinct if it is inside and contiguous with a rectangle. Hence:

$$\mu_{distinct}(c_i) = \max_{g_j \in G_{REC}} [\mu_{contig}(c_i, g_j) \cap \mu_{inside}(c_i, g_j)] \quad (3)$$

where $\mu_{contig}(c_i, g_j)$ and $\mu_{inside}(c_i, g_j)$ are the fuzzy memberships for contiguity and inside for the candidate as defined in Section 3.2.

Each scoring section has a definition in terms of unary and non-specific binary metrics. The candidates for each scoring section have the appropriate metrics calculated which are aggregated to give the candidates similarity measure. The unary metrics are aggregated using a weighted generalised mean. The relevant non-specific binary metrics are combined with these results using the Yager intersection (AND) function [31] to give the candidate's final result.

$$\mu_{unary}(c_i) = \bigcap_{k=0}^{n}\left(\sum_{i=1}^{p} w_{pk} m_{pk}^{\alpha_k}\right)^{1/\alpha_\kappa} \quad (4)$$

where m_{pk} is the pth metric for scoring section k with the weight, w_{pk}. Candidates below a working threshold are discarded.

The non-specific binary metrics used for each scoring section are listed in Table 1.

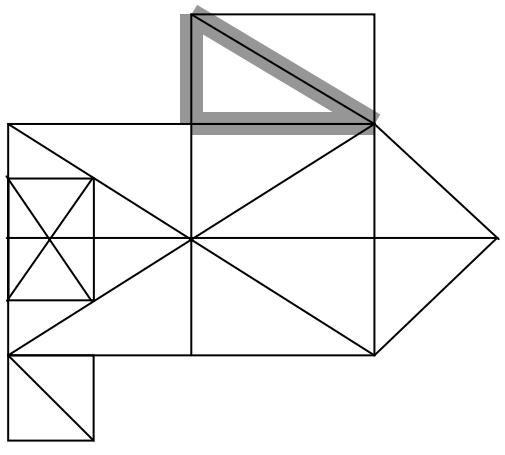


Fig. 5 *Example of an indistinct feature*

Table 1: Metrics used to reduce candidates

Scoring section 2	Largest
Scoring section 9	Extreme point (vertical), distinct (vertical)
Scoring section 13	Extreme point (horizontal), distinct (horizontal)
Scoring section 14	Extreme point (horizontal)
Scoring section 18	Extreme point (vertical)

3.2 Binary relations

The identification of a particular section is based on its relationship with other sections. This is a difficult task since any particular section can be misplaced, missing or repeated. Many distances are best considered relative to the size of the shape they relate to. However, we discovered that, in some extreme situations, this was either too harsh or too lenient [6], and so an absolute factor was also included. Many of the metrics are considered independently in the x and y axes. If the highlighted rectangle in Fig. 6 is considered then it is clear that its position is correct in the horizontal axis but incorrect in the vertical axis.

A summary of the metric types is given in Table 2.

The relative size of two features is based upon the distance between the extreme points in a given axis. However, if two features are contiguous then the size of the sides touching are of most importance and so a second size metric (contiguous size) is defined. The size metric is calculated as shown in (5).

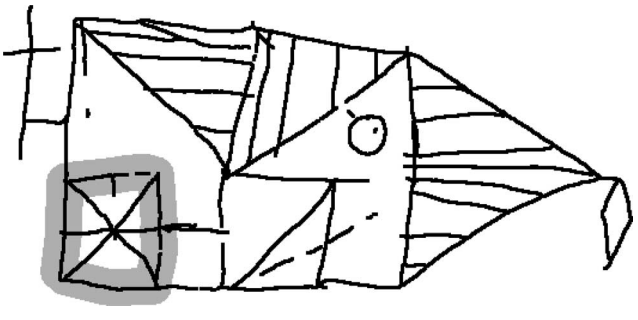


Fig. 6 *Section incorrectly positioned in one axis*

Table 2: Summary of binary metrics

Metric category	Summary
Relative size and position	Measure of the size in a given plane and position
Contiguous and touching	Measure of two features touching along a length or at a point
Inside, outside and direction	Measure of whether a feature is inside, outside or in a given direction compared to another feature
Point outside, direction and distance	As above but considers a single point rather than the whole feature
Orientation	Measure of relative orientation between features

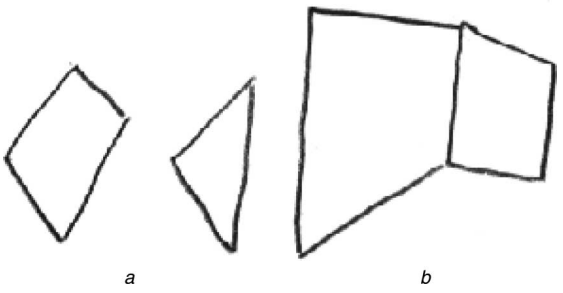


Fig. 7 *Examples of features of similar size*

a Vertically
b Contiguously

$$u_{size}(c_i, c_j) = u_h\left(\frac{s_i}{s_j}\right) \tag{5}$$

where s_i and s_j are the sizes of candidate c_i and c_j respectively in the given axis and differ for relative size and continuous size. This results in two metrics; Fig. 7*a* shows two features with a similar height (vertical size) while Fig. 7*b* shows two features with similar contiguous size.

The relative position is based upon the location of a shape's corner, with respect to the candidate's side. This allows an incorrectly sized shape to possibly score correctly for one edge and in effect produces an aggregation method of spatial relations [32]. This is described in (6)

$$\mu_{position}(c_i, c_j) = \mu\left(\frac{d}{l}\right) \tag{6}$$

where d is the absolute distance from the start of the reference figure, l is the length of the feature, as shown in Fig. 8 and is the result expressed as a fuzzy membership by using the full trapezoidal function, μ [30].

The contiguous nature of some scoring shapes is an important perceptual feature and is reflected by the contiguous metric. The grouping process produces a given number of comparison points along the side of each shape. The perpendicular distance to the other feature is calculated at each comparison point and aggregated to give a perpendicular closeness thus

$$\mu_{PerpClose}(c_i, c_j) = \left(\sum_{dp_n \in DP} \frac{1}{|DP|}\mu_r(dp_n, h)^{\alpha}\right)^{\frac{1}{\alpha}} \tag{7}$$

where μ_r returns a combined fuzzy membership of both absolute and relative measures; all the perpendicular distances of the grouping points, dp_n, on the side in the given axis are in the set DP and h is the smallest height of both candidates.

The two close sides must also be overlapped in a parallel direction, hence the contiguous membership is given in (8)

$$\mu_{contig}(c_i, c_j) = \mu_{PerpClose}(c_i, c_j) \cap \mu_r(o, l) \tag{8}$$

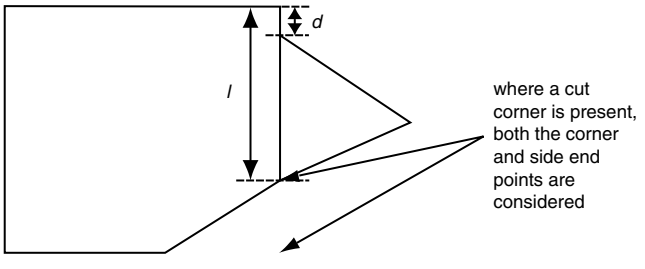


Fig. 8 *Example of relative position*

where o is the absolute distance that the two sides overlap and l is the shortest length of the two candidates. Hence, Fig. 9*a* shows an example with perfect contiguity, while Fig. 9*b* and Fig. 9*c* have a lower contiguous membership.

The touching metric describes a point contact at a corner of a shape to a side on the other feature. The touching metric for candidate c_i side ends on c_j is:

$$\mu_{stouch}(c_i, c_j) = \max_{dpp_k \in DPP, dpl_k \in DPL}[\mu_r(dpp_k, h) \cap \mu_r(dpl_k, l)] \quad (9)$$

where DPP and DPL are the sets of all the perpendicular and parallel distances respectively from all side end points of c_i to all the sides of c_j. h and l are the smallest height and length of the candidates respectively. Hence, the touch metric is given in (10)

$$\mu_{touch}(c_i, c_j) = \max[\mu_{stouch}(c_i, c_j), \mu_{stouch}(c_j, c_i)] \quad (10)$$

Figure 10 shows some examples of touch metric; Fig. 10*a* is perfect while Figs. 10*b* and 10*c* have a lower membership but are still significant.

Features can be inside or outside another feature. The general outside metric of a polygon can be expressed as in (11)

$$\mu_{outside}(c_i, c_j) = \max[\mu_{oside}(c_i, c_j, 1) \ldots, \mu_{oside}(c_i, c_j, n)] \quad (11)$$

where there are n sides to the polygon, and $\mu_{oside}(c_i, c_j, m)$ is defined in (12)

$$\mu_{oside}(c_i, c_j, m) = \bigcap_{pd_i \in PDC_m} \begin{cases} \mu_h(pd_i) & if \quad pd_i > 0 \\ 1 & otherwise \end{cases} \quad (12)$$

where PDC_m is the set of the perpendicular distance of all the comparison points on c_i to the side m of c_j. Both of the lower squares shown in Fig. 11 are outside (and below); the left one having a lower outside membership.

The inside metric is the fuzzy complement of the outside metric: $\mu_{inside}(c_i, c_j) = 1 - \mu_{oside}(c_i, c_j)$. When considering a contiguous feature (such as rectangular scoring section 18) it

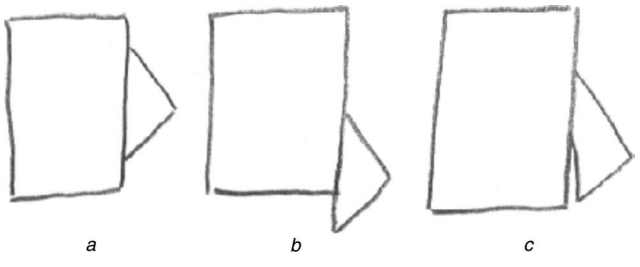


Fig. 9 *Various degrees of contiguity*

a Perfect contiguity
b,c Examples of lower configurous membership

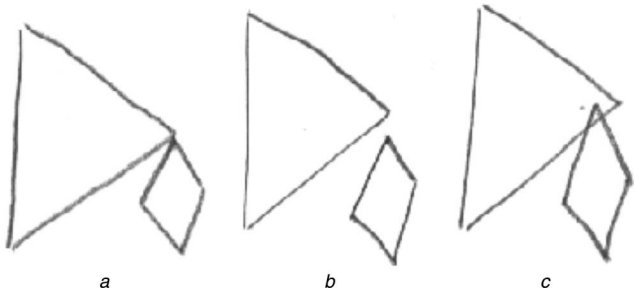


Fig. 10 *Various degrees of touch*

a Perfect
b,c Examples of lower membership

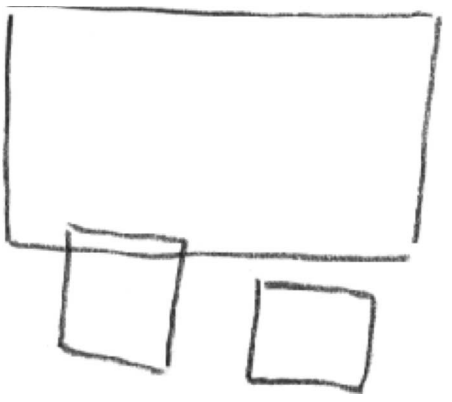


Fig. 11 *Squares outside*

is possible for a shape to be inside (or outside) in a given axis but not the other. If section 18 is considered then, while it is outside rectangle 2 in the vertical axis, it should be inside in the horizontal axis. This is achieved by expressing any external distance of the contiguous sides as a fuzzy metric.

A feature can be above, below, left or right of another shape. This is calculated in a similar manner to the outside metric except only the relevant side is considered. Symmetry is a perceptually significant feature and so a shape that has been misplaced with symmetry must be considered as more significant than one without.

Some features are in a given direction, inside or outside a reference shape which does not have a suitable side to use as reference. The triangle scoring section 13 should be to the right of triangle scoring section 9 (as well as below). With no contiguous sides it is not possible to use any sides as the reference and hence the extreme point is used. If the set of the distance from all the comparison points of c_i to the appropriate extreme side end point of c_j is DC, given that if the point falls between the two extreme points of c_j the distance is positive, then the outside point metric is given by (13)

$$\mu_{Pout}(c_i, c_j) = \sum_{dc_k \in DC} \left(\frac{1}{|DC|} \mu_d(dc_k)^{\alpha} \right)^{\frac{1}{\alpha}} \quad (13)$$

where

$$\mu_d(dc_k) = \begin{cases} \mu_r(dc_k, l) & if dc_k > 0 \\ 1 & otherwise \end{cases} \quad (14)$$

Features also have a relative orientation. The error between the closest angles of the sides of a shape (or axis for a diamond) is converted into a fuzzy membership. Hence, if the set A contained the differences in angles of all sides of c_i to all sides of c_j then:

$$\mu_{orien}(c_i, c_j) = \mu\left(\min_{a_i \in A}[a_i] \right) \quad (15)$$

This results in a metric that will still consider the features shown in Fig. 12 as having similar orientation.

Each scoring section has a number of metrics specified that define the relationship between each other scoring section. A weighted aggregation of all the relational metrics between two features is used to generate a final similarity measure for any two features. Hence, if the set of binary metrics that describe the relationship between scoring section i and j are described in BM_{ij}, each with an associated weight, w_k, then the similarity for the relationship between candidates $c_k \in SCORE_i$ and $c_l \in SCORE_j$ is calculated by (16)

Fig. 12 *Example of two features with the same orientation*

$$\mu_{binary}(c_k, c_l) = \left(\sum_{bm_k \in BM_{ij}} w_k bm_k^{\alpha} \right)^{\frac{1}{\alpha}} \quad (16)$$

This measure allows some metrics to be poor while still resulting in a high similarity.

3.3 Figure aggregation

To prevent combinatorial explosion we consider the figure broken into a number of subsections. The section candidates that make up the best subsection instances are combined to give the figure instances. The reduced implementation of the Osterrieth scoring system is covered by two subsections as shown highlighted in Figs. 13*a* and 13*b*.

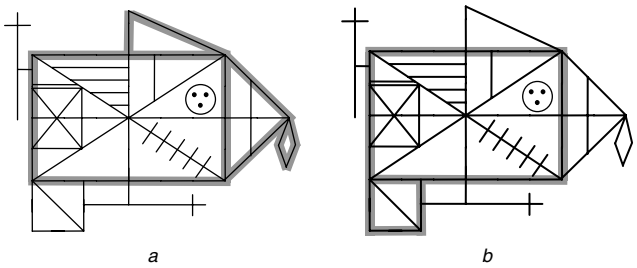
It is possible for a scoring section to be repeated and so the power set of candidates for each scoring section is generated, i.e. all possible combinations of the section's candidates are grouped into a number of different collections. This includes the empty set which is necessary since it is possible that all the candidates located are incorrect or that a section is not present. Inconsistent combinations of a section's candidates are discarded (inconsistent is defined as overlapping instances). The cartesian product of each member in these power sets generates a set of section instances that are used for a given subsection instance, denoted as SC_j for subsection instance j. Each scoring section candidate has a similarity membership calculated when considering it within the set of all the scoring section candidates for that subsection. This is denoted $\mu_{score}(c_j, SC_i)$ and is defined as

$$\mu_{score}(c_j, SC_i) = \bigcup_{c_k \in SC_i, c_j \neq c_k} w_j \mu_{binary}(c_j, c_k) \quad (17)$$

where w_k is the weight for the relationship between the scoring section of c_j and c_k such that $\sum w_a = 1$ for all the scoring sections within the figure subsection.

This results in the fuzzy OR function and so the similarity is of a given section is most influenced by the best relationship with any other feature. However, a Yager function [31] was used which still enabled some effect from other features.

The similarity for a given subsection instance is the aggregation of the binary similarity measure and is given by



a *b*

Fig. 13 *Osterrieth scoring system covered by two subsections*

a Subsection A
b Subsection B

$$\mu'_{sub}(SC_i) = \bigcap_{c_l \in SC_i} w_l . \mu_{score}(c_j, SC_i) \quad (18)$$

where w_l is the associated weight in the range [0,1]. Hence, a subsection's similarity is heavily influenced by the poorest section's similarity. However, the use of a Yager intersection (AND) enabled some influence by all the sections. Note: $\mu'_{sub}(SC_i)$ differs from $\mu_{sub}(SC_i)$, the final subsection similarity, by a coverage factor which is included to prevent the inclusion of trivial instances. This takes the form of a multiplying factor which is the ratio of scoring sections used compared to the number of scoring sections located.

A suitable instance from each subsection is combined to generate a figure instance. The union of the section candidates in each subsection instance is used to form the set of section candidates for the figure instance, such that $FC_i = SC_a \cup SC_b$. However, where subsections contain a common scoring section, only consistent subsection instances can be combined to generate a figure instance; i.e. the same section candidates for the common section must be used in all subsections. The similarity grade for a figure instance is calculated in the same manner as a subsection instance and so

$$\mu_{fig}(FC_i) = \bigcap_{c_j \in FC_i} w_j . \mu_{score}(c_j, FC_i) \quad (19)$$

where $\mu_{score}(c_j, FC_i)$ is the similarity measure for candidate c_j given the set of candidates FC_i. w_k is the associated weight.

Again, in a similar manner to the subsection calculation, a factor is applied to a figure's grade that indicates the number of sections used compared to the number found. Also, figure instances below a threshold from the highest figure instance were discarded.

As stated, a poor resultant grade does not necessarily indicate an incorrect result; however, every line of a figure should be identified. In a completed system, a quantity of unidentified figures can be used to indicate a suspicious interpretation and verify the grading.

4 Results

We tested the system using 31 random RCF drawings produced by children attending the Institute of Child Health, London, displaying a typical spread seen by the Neuropsychological Unit. Some 16 drawings were produced by copying the figure and 15 from recall. In total 140 scoring sections were to be identified and located by the system [6].

The location, rating and grading processes performed well, locating all but one scoring section and coinciding closely with subjective grades generated by independent raters [6]. After grouping, an average of 43 triangles, 77 rectangles and 42 diamonds were identified as candidates for scoring sections. Our use of unary metrics reduced

Table 3: The number of scoring section candidate located in the test data

	Scoring section	2	9	13	14	18
Number of candidates	Average	8.30	1.03	4.57	0.77	0.97
	Maximum	52	4	12	2	3
Number of combinations	Average	9.3	2.0	8.7	0.8	1.0
	Maximum	53	5	28	3	5

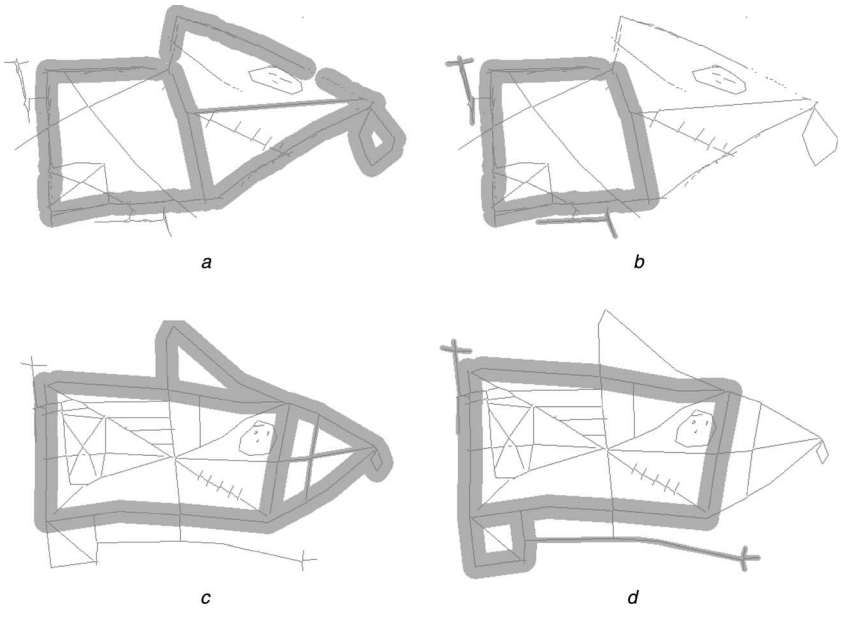


Fig. 14 *Examples of correctly identified subsections*
a,c Subsection A
b,d Subsection B

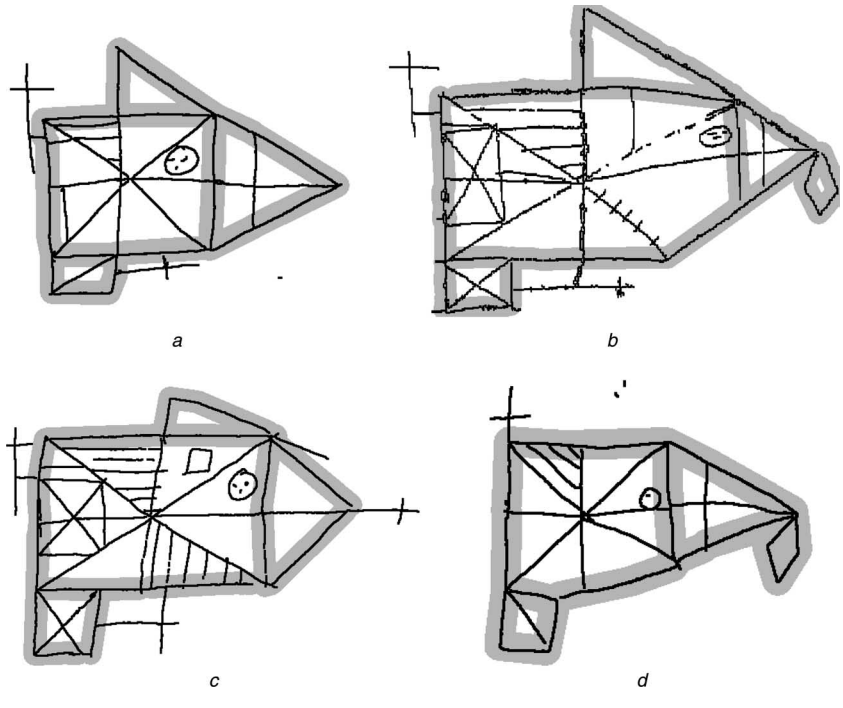


Fig. 15 *Example figures with correctly identified scoring sections*

these candidates by an average of 88% with no scoring sections discarded [7]. The non-specific metrics were then used to reduce the remaining candidates by an average of 56%, resulting in a total reduction in candidates of 96%. Again, no scoring section instances were discarded. The resultant number of scoring section candidates is given in Table 3.

The reduction in candidate numbers and a tendency for candidates for the same section being inconsistent resulted in the number of combinations for repeated scoring sections being acceptable.

The generation of the subsection instances performed well; an average of 419 instances of subsection A and 20 instances of subsection B were generated, causing no danger of combinatorial explosion. Of the 31 test figures all but three generated the correct instance for the data generated by the previous processing stages. A selection of correctly located subsections is shown in Fig. 14 with the identified features being highlighted.

Three subsections did not identify the correct scoring section instances. However, these were from figures that are distorted to such a degree that it is impractical to locate them automatically.

As we only describe a partial implementation here, the results need to be interpreted with some care. Using a suitable threshold to define a unique solution (i.e. where the similarity measure of the second highest is less than 0.9 of the highest) in only eight of the 31 test figures were all the scoring sections correctly selected from the candidate shapes. A selection of correctly identified scoring sections is shown in Fig. 15. Development of a module to disambiguate incorrectly identified scoring sections would

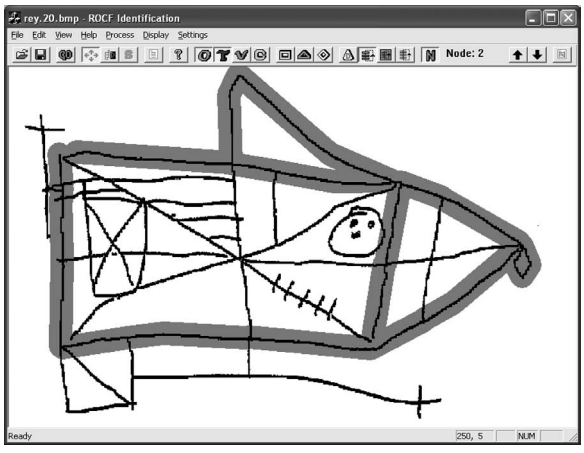


Fig. 16 *Semi-automated analysis system*

increase the success rate to approximately 75%, with the majority of the remaining figures being too difficult even for human scorers to interpret. While an initial extension of the system to handle such cases is straightforward, our experience with the complexity of interpreting highly distorted figures suggests that developing a robust module is a substantial challenge for future work.

As an intermediate solution to the problem, we have incorporated the processes described above into the semi-automated analysis system shown in Fig. 16. This system enables the user to observe the location and identification of the scoring structures interactively. Where a scoring section has been incorrectly identified, an alternate can be proposed by the system and selected by the user for inclusion in the marking process. Although this requires some user interaction, the subsequent grading of the scoring sections will still be fully automated.

This efficient and quantitative assessment of the subject's response to the RCF provides an objective alternative to the existing subjective manual scoring procedure. The semi-automated system employs a standard Windows interface with shortcut tool-bar buttons, which is simple to use and is familiar to even casual computer users. Initial use of the system has identified an increased confidence in the use of the RCF as a whole, as the process by which scoring sections are identified is more transparent and repeatable, removing much of the subjective input by the marker that has previously been the source of much concern. From this experience, it is confidently predicted that when complete, this system will reduce, significantly, the time taken to mark a patient's response and hence, lower the resource overheads in undertaking neurological assessment of this type.

It is important to stress, however, that the protocol for administering the RCF is unaffected. Normal practice regarding longitudinal, repeat and other testing schemes will be preserved allowing comparisons to be made in the traditional way.

5 Conclusions

In this paper we have presented a knowledge based system, employing fuzzy approximation techniques that can be used to locate the Osterrieth scoring sections of the Rey Complex Figure This neurological pen and paper test can produce a highly distorted line figure that poses an extremely challenging and novel problem for which no previous automated solution is available. Although at present a reduced set of scoring sections has been implemented to demonstrate the feasibility of the technique, we have shown how these can be incorporated into a semi-automated analysis system.

The system developed has demonstrated its ability to cope with clinical data in a practical manner that can be scaled to the full figure. Many of the techniques can also be applied to the process of generating the score of the sections to produce a completely automatic system. The technique therefore offers a robust and effective alternative to the existing subjective and resource intensive marking procedure.

Although the work described here has been applied specifically to the Osterrieth scoring sections of the Rey Complex Figure, there is no reason why the technique cannot be adapted to other neuropsychological tests of similar form. Specific examples include the Bender-Gestalt test [33] (which comprises a number of copying tasks, each a combination of two or more basic geometric shapes), the clock drawing test to measure executive cognitive dysfunction [34] and the Taylor Complex Figure [35], an alternative to the RCF. Considerable adaptation of the techniques reported in this paper is necessary to implement these tests, but the same benefits of a resource efficient, quantitative assessment is achievable.

6 Acknowledgment

We are extremely grateful to Dr Elizabeth Isaacs and her colleagues at the Institute of Child Health, London, for access to example patient responses to the RCF, and their contribution to this work.

7 References

1 Rey, A., and Osterrieth, P.: 'Translations of excerpts from Rey's 'Psychological examination of traumatic encephalopathy' and Osterrieth's 'The complex figure test'', *The Clin. Neuropsychol*., 1993, **7**, pp. 2–21
2 Meyers, J., and Meyers, K.: 'Rey complex figure test and recognition trial' (Psychological Assessment Resources, Odessa, Florida, 1995)
3 Tupler, L., Welsh, K., Asare-Aboagye, Y., and Dawson, D.: 'Reliability of the Rey-Osterrieth complex figure in use with memory - impaired patients', *J. Clin. and Exp. Neuropsyc*., 1995, **17**, (4), pp. 566–579
4 Bennett-Levy, J.: 'Determinants of performance on the Rey Osterrieth complex figure test: An analysis and a new technique for a single-case assessment', *Brit. J. Clin. Psychol*., 1984, **23**, pp. 109–119
5 Fastenau, P., Bennett, J., and Denburg, N.: 'Application of psychometric standards to scoring system evaluation: Is 'new' necessarily 'improved'?', *J. Clin. Exp. Neuropsyc*., 1996, **18**, (3), pp. 462–472
6 Canham, R., Smith, S., and Tyrrell, A.: 'Recognition of severely distorted geometric shapes from within a complex figure', *Pattern Anal. Appl*., 2000, **3**, (4) pp. 335–347
7 Canham, R., Smith, S., and Tyrrell, A.: 'Automated scoring of a neuropsychological test: The Rey Osterrieth complex figure', *Euromicro 2000*, 2000, **2**, pp. 406–413
8 Fairhurst, M.C., Smith, S.L., and Mitchel, J.: 'Automated image analysis in visuo-motor testing for the specification of an integrated evaluative and therapy support tool for rehabilitation', *IEEE Trans. Rehabil. Eng*., 1995, **3**, (1), pp. 103–111
9 Smith, S.L., and Hiller, D.: 'Image analysis of neuropsychological test responses', *Proc. SPIE*, 1996, **2710**, pp. 904–915
10 Guest, R.M., Fairhurst, M.C., and Potter, J.M.: 'Automated extraction of image segments from clinically diagnostic hand-drawn geometric shapes'. Proc. 26th EUROMICRO Conf. (EUROMICRO'00), Maastricht, The Netherlands, 2000, **2**, pp. 440–447
11 Guest, R.M., Fairhurst, M.C., and Potter, J.M.: 'Automatic classification of hand drawn geometric shapes using constructional sequence analysis'. Proc. Seventh Int. Conf. on Document Analysis and Recognition, IEEE Computer Soc, Los Alamitos, 2003, Vol. I, pp. 990–994
12 Bunke, H., and Messmer, B.: 'Similarity measures for structured representations'. Proc. of the First Euro. Workshop on Topics in Case Based Reasoning, 1994, pp. 106–118
13 Pasternak, B.: 'Processing imprecise and structually distorted line drawings by an adaptable drawing interpretation kernel'. Int. workshop on document analysis systems, 1994, pp. 318–337
14 Gross, M.: 'Recognizing and interpreting diagrams in design'. Proc. of the Workshop on Advanced Visual Interfaces, 1994, pp. 88–94
15 Havens, W., and Mackworth, A.: 'Schemata-based understanding of hand-drawn sketch maps'. Proc. 3rd Biennial Conf. of the Canadian

Society for Computational Studies of Intelligence, 1980, **45**, pp. 172–178
16 Goodson, K., and Lewis, P.: 'A knowledge based line recognition system', *Pattern Recognit. Lett.*, 1990, **11**, pp. 295–304
17 Hori, O., Shimotsuji, S., Hoshino, F., and Ishii, T.: 'Probabilistic relaxation method for line drawing interpretation'. 11th IAPR Int. Conf., 1992, **2**, (2), pp. 1158–1161
18 Partridge, D., and Hussain, K.: 'Knowledge based information systems' (McGraw - Hill, London, 1994)
19 Eshera, M., and Fu, K.: 'A graph distance measure for image analysis', *IEEE Trans. Syst. Man Cybern.*, 1984, **14**, (3), pp. 398–408
20 Kim, D., and Nevatia, R.: 'Recognition and localization of generic objects for indoor navigation using functionality', *Image Vis. Comput.*, 1998, **16**, pp. 729–743
21 Sandakly, F., and Giraudon, G.: '3d scene interpretation for a mobile robot', *Robot. Auton. Syst.*, 1997, **21**, pp. 399–414
22 Mees, W., and Acheroy, M.: 'Automated interpretation of aerial photographs', *Intell. Eng. Syst. Through Artif. Neural Netw.*, 1995, **15**, pp. 459–465
23 Lin, Y., Dou, J., and Zhang, E.: 'Edge expression based on tree structures', *Pattern Recognit.*, 1992, **25**, (5), pp. 507–517
24 Ralescu, A., and Shanahan, J.: 'Fuzzy perceptual grouping in image understanding'. IEEE/IFES '95, 1995, pp. 20–24
25 Sarkar, S., and Boyrt, K.: 'Using perceptual inference networks to manage vision processes', *Comput. Vis. Image Underst.*, 1995, **62**, (1), pp. 27–46
26 Keller, J., and Wang, X.: 'Comparison of spatial relation definitions in computer vision'. Proc. ISUMA/NAFIPS '95, 1995, pp. 679–684
27 Krishnapuram, R., and Ma, Y.: 'Quantitative analysis of properties and spatial relations of fuzzy image regions', *IEEE Trans. Fuzzy Syst.*, 1993, **1**, (3), pp. 222–233
28 Hilgard, E., Atkinson, R., and Atkinson, R.: 'Introduction to Psychology' (Harcourt Brace Jovanovich Inc, 1975 6th ed.)
29 Sarkar, S., and Boyer, K.: 'Perceptual organization in computer vision: A review and a proposal for a classification structure', *IEEE Trans. Syst. Man Cybern.*, 1993, **23**, (2), pp. 382–399
30 Lowen, R.: 'Mathematics and fuzziness' in 'Fuzzy set theory and applications' (Reidel Publishing, 1985)
31 Yager, R.: 'On a general class of fuzzy connectives', *Fuzzy Sets Syst.*, 1980, **4**, pp. 235–242
32 Keller, J., and Wang, X.: 'Learning spatial relations in computer vision'. Proc. 5th IEEE Int. Conf. on Fuzzy Systems, 1996, Vol. 1, pp. 118–124
33 Hutt, M.L.: 'The Hutt adaptation of the Bender–Gestalt test' (Grune & Stratton, New York, 1977 3rd edn.)
34 Juby, A., Tench, S., and Baker, V.: 'The value of clock drawing in identifying executive cognitive dysfunction in people with a normal mini-mental state examination score', *Can. Med. Assoc. J.*, 2002, **167**, (8), pp. 859–864
35 Taylor, L.B.: 'Psychological assessment of neurosurgical patients' in Rasmussen, T., and Marino, R. (Eds.): 'Functional neurosurgery' (Raven Press, New York, 1979)