Combining Integral and Separable Subspaces

Mikael Johannesson (mikael.johannesson@ida.his.se)
Department of Computer Science, University of Skövde, Sweden and
Lund University Cognitive Science, Lund University, Lund, Sweden

Abstract

It is well known that pairs of dimensions that are processed holistically - *integral dimensions* - normally combine with a Euclidean metric, whereas pairs of dimensions that are processed analytically - *separable dimensions* - most often combine with a city-block metric. This paper extends earlier research regarding information integration in that it deals with complex stimuli consisting of both dimensional pairs previously identified as holistic, and dimensional pairs previously identified as analytical. The general pattern identified is that information integration can be more accurately described with a rule taking aspects of stimuli into consideration compared to a traditional rule. For example, it appears that combinations of analytical and holistic stimuli, are better described by treating the different subspaces individually and then combining these with addition, compared to any single Minkowskian rule, and much better compared to any of the Minkowskian rules traditionally used (i.e. the city-block-, the Euclidean or the dominance-metrics).

Introduction

Suppose we have objects that differ on several dimensions – how is (dis-) similarity of such objects related to (dis-) similarity on each of the dimensions? Since Attneave (1950) raised essentially this question, much research efforts have been focused on the applicability of different combination rules. The most commonly investigated rules, or metrics, for describing distances in a multidimensional space have been instances of the generalised Minkowski metric (Eq. 1).

$$(1) \quad d(i,j) = \left\{ \sum_{k=1}^{n} | x_{ik} - x_{jk} |^{r} \right\}^{1/r} \quad ; \; r \geq 1$$

where $d(i,j)$ is the distance between object i and j, x_{ik} refers to the position of object i on the kth axis and n is the number of constituting dimensions.

Three extreme cases can be identified: r = 1; *the city-block metric* - The distance is the sum of the absolute differences for each of the underlying dimensions; r = 2: *the Euclidean metric* - The distance corresponds to the square root of the sum of the squared differences for each of the underlying dimensions; and r = ∞ : *the dominance metric* - The distance between two objects is a function of the distance for the separate dimension that differ the most.

When it comes to cognitive modelling, the city-block and, especially, the Euclidean metrics are the most common. However, it is well established that some pairs of dimensions combine, with Garners (1974) terminology, to form *integral dimensions* and others to form *separable dimensions* (see e.g. Garner, 1974, 1977). An integral pair typically is processed as holistic, unanalysable, directly and effortlessly by subjects and that the constituent dimensions combine so as to conform to a Euclidean metric; pairs of hue, saturation or brightness of colour (see e.g. Hyman & Well, 1967; Kemler Nelson, 1993) and the auditory dimensions of pitch and loudness (Kemler Nelson, 1993) typically do this. The corresponding description for a separable pair is that the constituent dimensions are processed independently by subjects and that they combine so as to conform to a city-block metric, e.g. size and reflectance of squares (Attneave, 1950).

Now, the combination rules for integral and separable dimensions are well investigated for dimensional *pairs*. But, what about more complex combinations? How do we integrate information when both integral and separable pairs are involved? Adequate descriptions of information integration behaviour is not only important from a theoretical perspective, but also from a more practical and pragmatic machine learning perspective.

Simple parallelograms varying in saturation, brightness, height and tilt could serve as an example. Pairs of the dimensions of colour, i.e. of hue, brightness and saturation, are often used as prototypical examples of integral dimensions (see e.g. Hyman & Well, 1967; Kemler Nelson, 1993). Perception of variation in saturation and brightness on a single colour patch have in previous studies (e.g. Hyman & Well, 1967; 1968) been shown to be better described using the Euclidean compared to the city-block metric. The height- (size-) and tilt- dimensions of parallelograms is an example of separable dimensions (Tversky & Gati, 1982). Tversky and Gati found such pairs to be better described using the city-block metric compared to the Euclidean.

How, then, could subjects' phenomenological (dis-) similarity between parallelograms varying in height, tilt, saturation and brightness be described? With reference to the different metric properties of the underlying pairs of dimensions, it makes sense to divide the stimuli space into two separate subspaces - one describing the aspects of shape of the stimuli (i.e. height and tilt) - *the shape space* - and one the colour aspects (i.e. saturation and brightness) - *the colour space*. In this case it could be that two different metrics should be applied: the city-block metric for the shape space and the Euclidean metric for the colour space. For combining the separate subspaces into a holistic measure, simple addition could be expected, with reference to that the pair of subspaces better fit the description of separability compared to integrality.

In the remainder of this paper, combination rules, such that the same Minkowski-r (r) applies to the whole stimuli space, will be referred to as *homogenous* rules. Metrics such that one r, say r_1, applies to one subspace, and one r, say r_2, applies to another, and that the holistic measure is obtained by combining the sub-metrics separably, will in the following be referred to as *heterogeneous* rules or metrics[1].

Now, how could we determine which of reasonable alternatives is the best when we want to describe (dis-) similarity judgements of stimuli varying in height, tilt, saturation and brightness?

General Method

The method used by Dunn (1983) in order to investigate the relationship between dimensional integrality and the combination rule used in a dissimilarity judgement task, will be adopted in this paper. However, it will be generalised in order to deal with stimuli with more than two underlying dimensions.

The basic idea is to divide the set of dissimilarity ratings into unidimensional and bidimensional ratings, reduce them to distances between points in a predefined dimensional space and then determine the r that best predicts the bidimensional dissimilarities from the unidimensional ones. In order to reduce ratings to distances correspondence, interdimensional additivity, intradimensional subtractivity and linearity must be assumed (Dunn, 1983; see also Johannesson, 2001a).

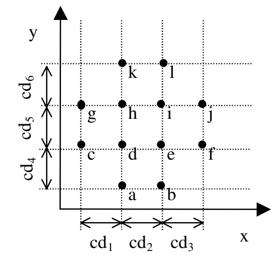

Figure 1: 12 stimuli and their component distances (cd1-cd6). (Based on Figure 1. in Dunn, 1983).

A first step is to decompose the unidimensional distances into component distances (see Figure 1 above). Further, under the assumption that the function relating dissimilarities to distances is linear, the dissimilarities between the stimuli in Figure 1 could, according to Dunn (1983), be expressed as

$$(2) \quad \delta(a,b) = \sum_{i=1}^{6} w_{iab} cd_i + A \;\; ; w_{iab} = 0 \text{ or } 1$$

where $\delta(a,b)$ is the perceived dissimilarity between object a and b, w_{iab} refers to the weight of the component distance cd_i, and A is an additive constant.

Eq. 2 specifies a multiple regression equation in which the weights define a set of dummy variables, the component distances form the regression coefficients and A is the additive constant.

Determining the Spatial Metric

Performing a multiple regression analysis on unidimensional dissimilarities provides an estimate of the component distances and the additive constant. From these, it is straightforward to estimate any "Minkowski dissimilarity". In order to determine the "best" describing metric for a particular subject, Dunn (1983) compared the mean observed and the mean predicted bidimensional dissimilarity using a certain value of r: overestimation of r lead to underestimation of the observed mean, whereas underestimation of r lead to overestimation of the observed mean.

Methodology Adopted

The methodology outlined by Dunn will be adopted with the exceptions as outlined below. Since the present paper aims to investigate whether the machine learning community could gain from using different Minkowski metrics for different subspaces rather than a single metric applied to the whole space, the various tests suggested by Dunn are not central here.

Experiment I

Subjects

14 students at the University of Skövde participated for a reward of cinema tickets roughly worth £11 or $17.

Stimuli

The stimuli were parallelograms varying in height (h; 4, 5 or 6 units of length), tilt (t; 40, 50 or 60 degrees), saturation (s; 40, 60 or 80% of maximum saturation) and brightness (b; 40, 60 or 80% of maximum brightness). The width and the hue of the parallelograms were held constant (4 units of length and 240 degrees of the colour circle, respectively).

In order to not exhaust the subjects, 20% of the possible 3240 pairs (i.e. 648) where chosen randomly. The order of the selected pairs were randomised.

[1] There are few examples of true integral dimensions in the literature (Grau & Kemler Nelson, 1988). This fact does not undermine the possible practical importance of heterogenous models, since perception of many dimensional pairs fall between the endpoints of a continuum of dimensional separability (Smith & Kilroy, 1979; Smith, 1980).

Procedure

The experimental sessions were performed individually in a quiet room with drawn curtains.

Each subject was first asked whether she/he had normal colour vision or not[2], and was then asked to follow the instructions given on the screen.

The experiment consisted of several phases:

- *Instruction phase*: Subjects were informed that they should judge dissimilarity between coloured parallelograms using a 20-graded scale.
- *Stimulus presentation*: Diminished versions of all stimuli were presented simultaneously in a randomised layout.
- *Training phase*: Subjects made dissimilarity judgements for ten pairs of coloured parallelograms varying in the same dimensions as the real stimulus material. The levels did not coincide with the levels of the real material.
- *Instruction phase*: An instruction phase as above was repeated. This time subjects were also informed that the judgement sessions would be divided into six parts with breaks between.
- *Stimulus presentation*: Subjects were again presented with the complete stimulus material.
- *Judgement phase*: The 648 stimulus pairs were presented in the same random order for all subjects. The experiment took about 2 hours.

All subjects reported they had normal colour vision.

Results and analysis

Table 1 below presents the average component distances (see Figure 1 above) per dimension, and the coefficient of determination for the collapsed data.

Table 1: Average component distances and R^2.

Avg_h	Avg_t	Avg_s	Avg_b	R^2
4.160	2.907	1.214	1.214	0.762

The average component distances, which could be interpreted as the relative saliency of each dimension (Dunn, 1983), differ between dimensions. Especially, the saturation and brightness dimensions have somewhat shorter component distances (are less weighted) compared to height and tilt. A possible explanation for this unequal weighting is that subjects perceived the variation in height and tilt as larger compared to the variation for the integral pair of saturation and brightness.

The coefficient of determination is not very large, indicating that a linear model misses to account for a considerable proportion of the variance of the data.

Determining the Spatial Metric When there are just two underlying dimensions it is obvious that distances should be estimated and evaluated for stimuli differing in two dimensions. However, as the number of underlying dimensions increases, so does the number of possibilities. In the present case, when four underlying dimensions were used, stimuli pairs differing in two or more dimensions were analysed.

Justifying the Measure of Error In order to possibly improve the process of determining the spatial metric, two alternative measures of error for a particular r were contrasted. One was in line with Dunn's method: deviation of the absolute difference between the mean observed dissimilarity and the mean predicted/estimated dissimilarity from the mean observed dissimilarity - in the following referred to as DEV. The other, referred to as the mean squared error (MSE), is defined as

$$(3) \quad MSE = \left(\sum_{\substack{a,b \\ a \neq b}} \left(\delta(a,b) - \tilde{\delta}(a,b) \right)^2 \right) / N$$

where $\delta(a,b)$ is the perceived, $\tilde{\delta}(a,b)$ is the predicted/estimated - dissimilarity between object *a* and *b*, and *N* is the number of stimuli pairs.

For each of the homogenous rules: city-block, Euclidean and dominance, and all non-ordered combinations of heterogeneous rules, where the subspaces where formed by the city-block, Euclidean or dominance metric[3], the distances between all non-ordered combinations of stimuli were calculated from physical descriptions of the stimuli. By regarding the distances as fictive dissimilarities, and by estimating the dissimilarities as described above for different rules, the errors according to DEV and MSE were calculated. The same subset and physical descriptions as used in the present experiment were analysed. Further, the estimated distances were scaled into a discrete scale ranging from 1 to 20. Since the underlying rule was known in each case, the two alternative measures of error could be evaluated against each other.

For the homogeneous models, both DEV and MSE suggested the same - and correct - underlying model. For the heterogeneous models MSE suggested the correct model in all cases. The use of DEV, however, was clearly systematically ambiguous. In all cases when

[2] In a pilot experiment preceeding this subjects performed a colour test in order to find out if they could discriminate between the colours that were to be used. Since all subjects in the pilot experiment reported the colour test to be simple a simple question was judged to be enough.

[3] Note that the heterogenous rule where both subspaces are formed by the city-block metric exactly corresponds to the city-block homogenous rule.

the underlying model could be described *as metric A applies to subspace 1 and metric B applies to subspace 2*, **both** the correct model and the model such that *metric B applies to subspace 1 and metric A applies to subspace 2*, were suggested. The explanation is that that the sum of absolute deviations for the two models necessarily is the same for a balanced set of stimuli.

In summary, based on this analysis, MSE appear to be the better measure for the purposes of this paper.

Spatial Metric Candidates for describing the individual subjects' data were evaluated using MSE as the measure of error. In addition to the rules used when evaluating the two error measures above, i.e.

- the homogenous rules: city-block, Euclidean and dominance - in the following referred to as *Hom cit*, *Hom euc* and *Hom dom*, respectively,
- all non-ordered combinations of heterogeneous rules, where each of the subspaces were formed by the city-block, Euclidean or dominance metric - in the following referred to as *Het citeuc*, *Het citdom*, *Het euccit*, *Hit euceuc*, *Hit eucdom*, *Hit domcit*, *Hit domeuc* and *Hit domdom*, respectively,

errors were calculated for values of Minkowski-r ranging in small discrete steps from r = 1.0 to r = 50.0 applied to the whole stimuli space (the homogenous model giving the lowest error will be referred to as *Hom opt*), the shape subspace and the colour subspace respectively. The heterogeneous model where the separately optimised r for the two-dimensional shape space is applied to "shape" and the separately optimised r for the two-dimensional colour space is applied to "colour", will be referred to as *Het sepHT-sepSB*. Finally, the combination of r:s, one for the shape subspace and one for the colour subspace, when optimised simultaneously with a heterogeneous rule - will be referred to as *Het simHTsimSB*.

Table 2: Models, r:s and errors for average data.

	R	Err
Het simHTsimSB	1.55;2.25	2.146
Het sepHTsepSB	1.55;2.2	2.146
Het euceuc	2;2	2.339
Hom opt	1.2	2.481
Het eucdom	2;50	2.601
Het euccit	2;1	2.894
Het domeuc	50;2	3.838
Het domcit	50;1	3.905
Het citdom	1;50	3.948
Het citeuc	1;2	4.194
Het domdom	50;50	4.313
Hom cit	1	5.907
Hom euc	2	7.805
Hom dom	50	16.644

The candidate models evaluated and the errors for the collapsed data are presented in Table 2 above.

A heterogeneous model combining a rule between the city-block and the Euclidean metrics[4] for the shape space, and a rule roughly corresponding to the Euclidean metric for the colour space (*Het simHTsimSB* and *Het sepHTsepSB*), gave a lower error than the best of the homogenous models (*Hom opt*), which had a r = 1.2, i.e. halfway between the city-block and the Euclidean metrics. This was true irrespectively of if the r:s were optimised separately or simultaneously. The optimal heterogeneous Minkowski-r:s found were lower for the shape space compared to the colour space. However, the r:s found were slightly different from the levels identified by previous research when two-dimensional stimuli have been used. The values were somewhat higher compared to what has been identified for these spaces before. It may be the case that the r-value goes up when the dimensionality increases. This speculation makes sense considering that we have limitations in terms of how many dimensions we can process simultaneously, and that larger values of r corresponds to focusing more on the dimension where the stimuli-pair at hand differ the most.

The common homogenous Euclidean rule (*Hom euc*) gave a substantially worse error than both the best heterogeneous rule and the best homogenous rule. However, the somewhat unequal weightings of the dimensions defining the two subspaces (see above) probably causes the peculiarity that *Het euccit* produces an error lower than that for *Het citeuc*. The fact that there are differences in weighting indicate that there are differences in salience between dimensions.

In summary, a heterogeneous rule or model seems to describe the data better compared to a homogenous one.

Errors and r:s were calculated also for the heterogeneous rules combining the "odd", or counterintuitive, subspaces h/s and t/b on one hand and h/b and t/s on the other. The heterogeneous models with the lowest errors for the average data for each of the three subspace divisions are presented in Table 3.

Table 3: Different subspace divisions and their errors.

Subspace division	Model	Err
height/tilt; sat./bri.	Het simHTsimSB	2.146
height/sat.; tilt/bri.	Het simHSsimTB	3.030
height/bri.; tilt/sat.	Het simHBsimTS	2.861

The errors for the heterogeneous models for the "odd" subspace divisions are considerably larger compared to the error for the original division. For the

[4] Note, however, that the Minkowsky-r of a rule giving distances halfway between the city-block and the Euclidean metric is not the intuitive 1.5, but rather approximately 1.2.

individual data, the corresponding difference was true for 8 out of 12 cases with at least one r differing from 1.0. This difference indicate that the intuitive division into subspaces of shape and colour makes sense.

Experiment II

In Experiment II, the heterogeneous r:s found were larger than what has been found in earlier research. A reasonable question is if the element of non-separability together with the increased dimensionality causes such effects. A second experiment was conducted in order to investigate if integrality (non-separability) could be eliminated as an explanation or not. Contrary to Experiment I, the underlying dimensions in the present experiment are purely separable.

Subjects

12 students (the majority were undergraduates) at the University of Skövde participated for a reward of cinema tickets roughly worth £11 or $17.

Stimuli

The stimuli varied in four dimensions, height (h), tilt (t), width of a stripe parallel to the horizontal axes (st) and brightness (b) of a parallelogram. These dimensions differ from the ones used in Experiment I above in some crucial aspects. One is that they do not form intuitive subspaces. Another is that all possible pairs of dimensions match the description of separable dimensions.

Each dimension varied in three levels, h: (4, 5 or 6 units of length), t: (40, 50 or 60 degrees), st: (1, 2 or 3 units of width) and b: (40, 60 or 80% of maximum brightness). The width, hue and saturation were held constant (4 units of length, 240 degrees and 60% of maximum saturation, respectively).

The same pairs (w.r.t. the numbers of the stimuli), and order between pairs as in Experiment I were used.

Procedure

The experiment was conducted as Experiment I above.

Results

The average component distances for the collapsed data in Experiment II (Table 4 below), are not perfectly equal, especially the brightness dimension is weighted less compared to the others.

Table 4: Average component distances and R^2.

Avg_h	Avg_t	Avg_st	Avg_b	R^2
2.089	2.381	1.530	0.625	0.541

The coefficient of determination is very low, hence a general linear model does not apply well.

Spatial Metric The same candidate models as evaluated in Experiment I were evaluated. The resulting errors for the collapsed data are presented in Table 5.

Table 5: Models, r:s and errors for average data.

	r	Err
Het simHTsimSTB	1.1; 1	3.483
Hom opt	1	3.523
Hom cit	1	3.523
Het sepHTsepSTB	1.6; 1	4.010
Het citeuc	1; 2	4.213
Het euccit	2;1	4.434
Het citdom	1; 50	4.638
Het euceuc	2; 2	5.573
Het domcit	50; 1	5.885
Het eucdom	2; 50	6.185
Het domeuc	50; 2	7.234
Het domdom	50; 50	7.935
Hom euc	2	11.333
Hom dom	50	17.219

It is clear that the best rule, of the ones tested for, for describing the collapsed data in Experiment II is close to a city-block rule (*Het simHTsimSTB* (r=1.1;1), *Hom opt* (r=1) and *Hom cit*). It is not, in this special case, possible to view this as supporting either of homogenous or heterogeneous models since the city-block metric is the sum of the differences for the constituting dimensions. Therefore, there is no difference between a homogenous city-block rule and a heterogeneous rule where city-block rules are used within all subspaces.

As opposed to experiment I, the Minkowski-r values (for the best models) did not increase in magnitude with increased dimensionality.

The heterogeneous models with the lowest errors for the average data for each of the three subspace divisions are presented in Table 6. As, for the collapsed data, the optimal "heterogeneous" rule for the "original" subspace division was close to the city-block metric for both subspaces, this was necessarily the case also for the "odd" subspace divisions.

Table 6: Different subspace divisions and their errors.

Subspace division	Model	Err
height/tilt; str./bri.	Het simHTsimSTB	3.483
height/str.; tilt/bri.	Het simHSTsimTB	3.523
height/bri.; tilt/str.	Het simHBsimTST	3.523

General Discussion

The aim of this paper is to investigate and communicate the idea that division of features/dimensions of objects into separate subspaces - when applicable - possibly could increase descriptive power.

Experiment I involved pairs of dimensions previously found to be combined best by the city-block and the Euclidean metric, respectively. The Euclidean rule turned out to badly describe the data. Instead, a heterogeneous rule combining the two subspaces formed by the intuitive division, was found to provide the best description. The r:s for the two subspaces found in this experiment rhymes with previous research in that they really possess different metric properties and that the r for saturation/brightness was higher than for height/tilt. However, both r:s found were somewhat larger compared previous findings for the separate two-dimensional subspaces. The dimensions involved in Experiment I were all expected to be pairwise separable. Also in the four-dimensional case, the best describing metric turned out to be the city-block rule.

The idea presented received support in that the general pattern identified from the experiments is that phenomenological dissimilarity can be more accurately described with a heterogeneous rule taking aspects of the stimuli into consideration, compared to a homogenous Minkowski-metric.

There are a number of open questions. One relevant issue is how the subspaces themselves should be combined. In this paper, only one of many possible ways was investigated. Another question concerns the magnitudes of the r:s identified. Since the r:s estimated in Experiment II were not larger compared to what could be expected for pairwise combinations of the constituent dimensions, it is apparent that the increase in magnitude of r:s as found in Experiment I, is not generalisable to all complex stimuli. However, it is in the developmental literature well documented that the separability changes with experience (see e.g. Smith, 1980), with the direction from integrality to separability. This pattern also apply to short term learning (Johannesson, 2001b). A possible reason for the relatively large r:s in Experiment I could thus be that stimuli with contents of integrality are harder to "learn" than stimuli composed by separable dimensions. If so, the r:s could possibly stabilise at a lower magnitude for sufficiently experienced subjects. If not, it could simply be that the specific metric properties associated with integral/separable dimensions only are true in the context of single pairs of dimensions, i.e. depending on if they are combined or not. An interesting set of stimuli that could be used in order to explore this (and others) issue further is multimodal stimuli composed of the pairwise integral dimensions of pitch/loudness and hue/saturation.

The results presented clearly motivates further research on the idea that information integration could be described as a combination of distances within different subspaces. More research on if, how and when information integration behaviour can be described in terms of combinations of subspaces may shed light on how we interact with the inherently high-dimensional real world. For example, Edelman & Intrator (1997) discuss the necessity of low dimensionality for learning in perceptual tasks - known as 'the curse of dimensionality'. However, even if we always use low-dimensional representations internally, even for cognition, if these representations involve more than two dimensions, cognitive science have interesting problems to solve.

Acknowledgements

This project have been financed by the Department of Computer Science at University of Skövde and by a grant to the Center of Learning Systems at University of Skövde from the Foundation for Knowledge and Competence Development (1507/97), Sweden. I wish to thank Professor Lars Niklasson, University of Skövde, and Professor Peter Gärdenfors, Lund University Cognitive Science, for support and comments.

References

Attneave, F. (1950). Dimensions of similarity. *American Journal of Psychology, 63*, 516-556.

Dunn, J. C. (1983). Spatial metrics of integral and separable dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 9 (2)*, 242-257.

Edelman, S. and Intrator, N. (1997). Learning as formation of low-dimensional representation spaces. In Shafto, M.G. and Langley, P., editors, *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, Erlbaum, Mahwah, NJ, 199-204.

Garner, W. R. (1974*). The processing of information and structure*, New York, Wiley.

Garner, W. R. (1977). The effect of absolute size on the separability of the dimensions of size and brightness. *Bulletin of the Psychonomic Society, 9 (5)*, 380-382.

Hyman, R. and Well, A. (1967). Judgments of similarity and spatial models. *Perception & Psychophysics, 2 (6).*

Hyman, R and Well, A. (1968). Perceptual separability and spatial models. *Perception & Psychophysics, 3*, 161- 165.

Johannesson, M. (2001a). The Problem of Combining Integral and Separable Dimensions. *Tech. rep. at the Department of Computer Science*. University of Skövde, Sweden., and *Lund University Cognitive Studies*. Lund University , Sweden.

Johannesson, M. (2001b). Toward Separability During Learning. *Submitted for publication*.

Kemler Nelson, D. G. (1993). Processing integral dimensions: The whole view. *Journal of Experimental Psychology: Human Perception and Performance, 19 (5)*, 1105 - 1113.

Tversky, A. and Gati, I. (1982). Similarity, separability, and the triangle inequality. *Psychological Review, 89 (2)*, 123-154.