2-D ANALOGUES OF ALLEN INTERVAL ALGEBRA FOR IMAGE ANALYSIS: TOWARDS JUSTIFICATION

SCOTT A. STARKS, DIMA IOURINSKI, and VLADIK KREINOVICH

NASA Pan-American Center for Earth and Environmental Studies
University of Texas, El Paso, TX 79968, USA, vladik@cs.utep.edu

Abstract

In reasoning about time and duration, researchers often use Allen's Interval Algebra. This algebra describes possible relations between 1-D intervals. An interval can precede the other one, follow the other one, start the other one, etc. This algebra describes the relationship between different intervals in terms of words from natural language. To give a natural language description of 2D images, it is desirable to develop a similar approach for describing the relationship between 2-D objects in a picture. In their recent papers, Jim Keller and his collaborators proposed a new approach based on a simulation of a "force" between these objects. In this paper, we show that their force formula is theoretically optimal.

Keywords
Image analysis, relation between images.

1 Introduction

1.1 How to describe relation between 2-D objects: informal description of the problem

In reasoning about time and duration, researchers often use Allen's Interval Algebra; see, e.g., [2–5]. This algebra describes possible relations between 1-D intervals. An interval can precede the other one, follow the other one, start the other one, etc. This algebra describes the relationship between different intervals in terms of words from natural language.

To give a natural language description of 2D images, it is desirable to develop a similar approach for describing the relationship between 2-D objects in a picture.

1.2 Force histogram method of describing the relation between 2-D objects

A new method of describing the relation between the 2-D objects, a method that seems to be in a very good accordance with the expert reasoning, was recently described in [6], [7], [9]. Namely, to describe the relation between the two sets A and B, this "force histogram" method proposed, crudely speaking, to do the following:

- take all possible pairs $a \in A$ and $b \in B$,
- compute, for each pair, the "force" whose direction is from a to b and whose value is r^{-2}, and then,
- add up these force vectors.

The orientation of the resulting vector $\vec{F}$ describe the intuitive understanding of where the sets are in relation to each other, and the length of this vector describes how close these sets are to each other.

1.3 This method is somewhat ad hoc

In the above description of the force method, the idea that we should:

- take all possible pairs compute, for each pair of points $a \in A$ and $b \in B$, the "force" whose direction is from a to b and whose value $f(r)$ depends on the distance between the points, and then,
- add up these force vectors

is very natural. What is not very natural is the exact form of the dependence on $f(r)$ on r. It is therefore desirable to consider all possible dependencies and come up with the optimal (best possible) function $f(r)$.

1.4 What we are planning to do

In order to solve this problem, we must describe, in precise terms, which functions are possible and which functions are the best. In this paper, we propose such a description, and show that this description leads to a new symmetry-based justification for the above formula (actually, for a slightly more general formula $r^{-\alpha}$).

This new justification is in line with the general symmetry-based approach which has been shown, in [8], to explain similar heuristic formulas in fuzzy, neural, genetic, and other approaches, and to explain a similar heuristic force formula in robotic control.

2 First Idea: Only Decreasing Functions $f(r)$ Make Sense

The function $f(r)$ describes the "weight" with which directions between different pairs (a, b) influence the resulting sum. Intuitively, when we decide which of the two objects is, say, to the left and which is to right, we pay more attention to close points and less attention to points which are far away from each other. Therefore, it is reasonable to require that the weight $f(r)$ should be larger for closer points and smaller for more distant points. In other words, it is reasonable to require that the function $f(r)$ be strictly decreasing.

3 Second Idea: We Must Choose a Family of Functions, Not a Single Function

Once we select a force function $f(r)$, for every two sets A and B, the orientation of the resulting vector $\vec{F}$ describe the intuitive understanding of where the sets are in relation to each other, and the length $\ell \stackrel{\text{def}}{=} |\vec{F}|$ of this vector describes how close these sets are to each other.

Our main goal is to describe the orientation, i.e., the typical angle between the two sets. We also get some information from the value of the "closeness" ℓ, but this information comes only from the comparison of different pairs of sets. If we simply say that the closeness between two sets A and B is, say, 1.5, we did not gain any knowledge about how close they are. If, however, for some other pair of sets A' and B', the closeness is 4.5, we can already make some meaningful conclusions: namely, we can say that A is closer to B than A' is close to B'. We can even say that A is three times closer to B than A' is close to B'.

From this viewpoint, the absolute values of F do not carry much information, what carries information is the relation between the values F corresponding to different pairs of sets.

In particular, if we multiply all the values of the original function $f(r)$ by a constant, i.e., if we replace the function $f(r)$ with a new function $\widetilde{f}(r) \stackrel{\text{def}}{=} C \cdot f(r)$ for some constant C, then all the values of the forces get multiplied by C and hence, all the resulting vectors $\vec{F}$ are multiplied by the same constant C. When we multiply a vector by a constant, it orientation does not change, and its length gets multiplied by the same constant.

Thus, if we replace the function $f(r)$ by a new function $C \cdot f(r)$, for every pair of sets, we get the exact same orientation as before, and for every two pairs, we get the exact same ratio between their "closenesses". So, intuitively, there is no big difference between using the original function $f(r)$ and the new function $C \cdot f(r)$.

Hence, we cannot select a unique function $f(r)$ and claim it to be the best, because for every function $f(r)$, the function $C \cdot f(r)$ leads to exactly the same results. In view of this, instead of formulating a problem of choosing the best force *function*, it is more natural to formulate a problem of choosing the best *family* $\{C \cdot f(r)\}_C$ of force functions.

4 Which Family Is the Best? We May Need Non-Numerical Optimality Criteria

Among all the families $\{C \cdot f(r)\}_C$, we want to choose the best one.

In mathematical optimization problems, numerical criteria are most frequently used, when to every alternative (in our case, to each family) we assign some value expressing its performance, and we choose an alternative (in our case, a family) for which this value is the largest. In our problem, as such a numerical criterion, we can select, e.g., the average approximation error A, measured as the mean square deviation between the orientation generated by the corresponding force method and the orientation marked by an expert.

However, it is not necessary to restrict ourselves to such numerical criteria only. For example, if we have several different families that have the same average approximation error A, we can choose between them the one that has the minimal computation time T. In this case, the actual criterion that we use to compare two families is not numerical, but more complicated: *a family F_1 is better than the family F_2 if and only if either $A(F_1) < A(F_2)$, or $A(F_1) = A(F_2)$ and $T(F_1) < T(F_2)$.* A criterion can be even more complicated. What a criterion *must* do is to allow us for every pair of families to tell whether the first family is better with respect to this criterion (we'll denote it by $F_1 \succ F_2$), or the second is better ($F_1 \prec F_2$) or these families have the same quality in the sense of this criterion (we'll denote it by $F_1 \sim F_2$). Of course, it is necessary to demand that these choices be consistent, e.g., if $F_1 \prec F_2$ and $F_2 \prec F_3$ then $F_1 \prec F_3$.

5 The Optimality Criterion Must Select a Unique Optimal Family

Another natural demand is that this criterion must choose a *unique* optimal family (i.e., a family that is better with respect to this criterion than any other family). The reason for this demand is very simple.

If a criterion does not choose any family at all, then it is of no use.

If several different families are "the best" according to this criterion, then we still have a problem to choose among those "best". Therefore, we need some additional criterion for that choice. For example, if several families turn out to have the same average approximation error, we can choose among them a family with the minimal computation time.

So what we actually do in this case is abandon that criterion for which there were several "best" families, and consider a new "composite" criterion instead: F_1 is better than F_2 according to this new criterion if either it was better according to the old criterion or according to the old criterion they had the same quality and F_1 is better than F_2 according to the additional criterion.

In other words, if a criterion does not allow us to choose a unique best family it means that this criterion is not final. We have to modify it until we come to a final criterion that will have that property.

6 The Optimality Criterion Must Be Scale-Invariant

The next natural condition that the criterion must satisfy is connected with the fact that the numerical value of the distance r depends on the choice of the unit for measuring distance.

If we replace the original unit of length by a new unit which is λ times larger (i.e., replace feet by meters), then numerical values change from r to $\widetilde{r} = r/\lambda$. How will the force function look in the new units? Let us assume that in the new units, the distance between the two points equals $\widetilde{r}$. Then, the same distance in the old units is equal to $r = \lambda \cdot \widetilde{r}$. Thus, the force between the two points is equal to $f(r) = f(\lambda \cdot \widetilde{})$. Thus, if we know the distance $\widetilde{r}$ in the new units, we can compute the corresponding force value as $\widetilde{f}(\widetilde{r})$, where $\widetilde{f}(z)$ denotes $f(\lambda \cdot z)$. So, the same force function which, in the old units, had the form $f(r)$, in the new units, has a new form $f(\lambda \cdot r)$.

Since this change is simply a change in a unit of length, it is reasonable to require that going from $f(r)$ from $f(\lambda \cdot r)$ should not change the *relative* quality of the force functions, i.e., if a family $\{C \cdot f(x)\}_C$ is better that the family $\{C \cdot g(r)\}_C$, then for every $\lambda > 0$, the family $\{C \cdot f(\lambda \cdot r)\}_C$ must be still better than the family $\{C \cdot g(\lambda \cdot r)\}_C$.

So, we arrive at the following definitions.

7 Definitions and the Main Result

Definition 1.

- *By a force function we mean a strictly decreasing function from non-negative real numbers to non-negative real numbers.*
- *By a family of functions we mean the family $\{C \cdot f(r)\}_C$, where $f(r)$ is a given force function and C runs over arbitrary positive real numbers.*
- *A pair of relations $(\prec, \sim)$ is called consistent [8] if it satisfies the following conditions:*
  (1) *if $a \prec b$ and $b \prec c$ then $a \prec c$;*
  (2) *$a \sim a$;*
  (3) *if $a \sim b$ then $b \sim a$;*
  (4) *if $a \sim b$ and $b \sim c$ then $a \sim c$;*
  (5) *if $a \prec b$ and $b \sim c$ then $a \prec c$;*
  (6) *if $a \sim b$ and $b \prec c$ then $a \prec c$;*
  (7) *if $a \prec b$, then $b \prec a$ or $a \sim b$ are impossible.*

Definition 2.

- *Assume a set A is given. Its elements will be called alternatives. By an optimality criterion we mean a consistent pair $(\prec, \sim)$ of relations on the set A of all alternatives. If $b \prec a$, we say that a is better than b; if $a \sim b$, we say that the alternatives a and b are equivalent with respect to this criterion.*
- *We say that an alternative a is optimal (or best) with respect to a criterion $(\prec, \sim)$ if for every other alternative b either $b \prec a$ or $a \sim b$.*
- *We say that a criterion is final if there exists an optimal alternative, and this optimal alternative is unique.*
- *Let $\lambda > 0$ be a real number. By the λ-rescaling $R_\lambda(f)$ of a function $f(r)$ we mean a function $(R_\lambda(f))(r) \stackrel{\text{def}}{=} f(\lambda \cdot r)$.*
- *By the λ-rescaling $R_\lambda(F)$ of a family F, we mean the set of the functions that are obtained from $f \in F$ by λ-rescaling.*

In this paper, we consider optimality criteria on the set $\mathcal{F}$ of all families.

Definition 3. *We say that an optimality criterion on F is scale-invariant if for every two families F and G and for every number $\lambda > 0$, the following two conditions are true:*

i) *if F is better than G in the sense of this criterion (i.e., $G \prec F$), then*

$$R_\lambda(G) \prec R_\lambda(F);$$

ii) *if F is equivalent to G in the sense of this criterion (i.e., $F \sim G$), then*

$$R_\lambda(F) \sim R_\lambda(G).$$

As we have already remarked, the demands that the optimality criterion is final and scale-invariant are quite reasonable. The only problem with them is that at first glance they may seem rather weak. However, they are not, as the following Theorem shows:

Theorem. *If a family F is optimal in the sense of some optimality criterion that is final and scale-invariant, then every force function $f(r)$ from this optimal family F which has the form $f(r) = A \cdot r^{-\alpha}$ for some real numbers A and α.*

Thus, our theorem says, in effect, that the force functions used by J. Keller (or, to be more precise, a slightly more general class of force functions) are indeed optimal.

8 Proof

1. Let us first prove that the optimal family $F_{\rm opt}$ is scale-invariant, i.e., that $R_\lambda(F_{\rm opt}) = F_{\rm opt}$ for all λ and λ.

Indeed, let $\lambda > 0$ be a positive real number. By definition of optimality, for every family $F \in \mathcal{F}$, we have $F \preceq F_{\rm opt}$. In particular, we can conclude that $R_{1/\lambda}(F) \preceq F_{\rm opt}$. From the scale-invariance of the optimality criterion, we can now conclude that $R_\lambda(R_{1/\lambda}(F)) \preceq R_\lambda(F_{\rm opt})$, i.e., that $F \preceq R_\lambda(F_{\rm opt})$. This is true for every family $F \in \mathcal{F}$ and therefore, the family $R_\lambda(F_{\rm opt})$ is optimal. But since the criterion is final, there is only one optimal family; hence, $R_\lambda(F_{\rm opt}) = F_{\rm opt}$. So, the optimal family is indeed scale-invariant.

2. Scale-invariance $R_\lambda(F_{\rm opt}) = F_{\rm opt}$ means, in particular, that for every function $f(r)$ from the optimal family $F_{\rm opt}$, we have $R_\lambda(f) \in F_{\rm opt}$. By definition, $R_\lambda(f)$ means $f(\lambda \cdot r)$, and $F_{\rm opt}$ consists of all the functions of the type $C \cdot f(r)$. Thus, we conclude that for every $\lambda > 0$, there exists a $C > 0$ (depending on λ) for which, for every r, we have

$$f(\lambda \cdot r) = C(\lambda) \cdot f(r). \qquad (1)$$

All decreasing solutions of this functional equations are known (see, e.g., [1]); these solutions are $f(r) = A \cdot r^{-\alpha}$. Thus, the theorem is proven.

For readers' convenience, let us describe how this can be proven when $f(r)$ is a differentiable function. In this case, $f(\lambda \cdot r)$ is a differentiable function as well, hence their ratio $C(\lambda) = f(\lambda \cdot r)/f(r)$ is also differentiable. Thus, both sides of the equality (1) are differentiable with respect to λ. Differentiating both sides relative to λ and substituting $\lambda = 1$, we get the following equation:

$$r \cdot \frac{df(r)}{dr} = C_0 \cdot f(r), \qquad (2)$$

where C_0 denotes the derivative of the function $C(\lambda)$ at $\lambda = 1$. Multiplying both sides of the equation (2) by $dr/(r \cdot f)$, we conclude that:

$$\frac{df}{f} = C_0 \cdot \frac{dr}{r}.$$

Integrating, we get $\ln(f) = C_0 \cdot \ln(r) + C_1$. Hence,

$$f = \exp(C \cdot \ln(r) + C_1) = \exp(C_1) \cdot r^{C_0},$$

i.e., the desired formula for $A = \exp(C_1)$ and $\alpha = -C_0$.

Acknowledgments

This work was supported in part by NASA grants NCC5-209 and NCC 2-1232, by the Air Force Office of Scientific Research grant number F49620-00-1-0365, by Grant No. W-00016 from the U.S.-Czech Science and Technology Joint Fund, and by Grant NSF 9710940 Mexico/Conacyt.

The authors are thankful to Jim Keller for useful discussions.

References

[1] J. Aczel, Lectures on functional equations and their applications, Academic Press, N.Y.–London, 1966.

[2] J. Allen, Maintaining knowledge about temporal intervals", Communications of the ACM, 1983, Vol. 26, pp. 832–843 (reprinted in [10]).

[3] J. Allen, Towards a general theory of action and time, Artificial Intelligence, 1984, Vol. 23, pp. 123–154.

[4] J. Allen and P. Hayes, A common-sense theory of time, Proceedings of the 9th International Joint Conference on Artificial Intelligence, Los Angeles, 1985, pp. 528–531.

[5] J. F. Allen and H. A. Kautz, A model of naive temporal reasoning, In: J. R. Hobbs and R. C. Moore (eds), Formal theories of the commonsense world, Ablex, Norwood, NJ, 1985, pp. 251–268.

[6] J. M. Keller and P. Matsakis, Aspects of hugh-level computer vision using fuzzy sets, Proceedings FUZZ-IEEE'99, Seoul, Korea, 1999, pp. 847–852.

[7] P. Matsakis and L. Wendling, A new way to represent the relative position between areal objects, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, Vol. 21, No. 7, pp. 634–643.

[8] H. T. Nguyen and V. Kreinovich, Applications of continuous mathematics to computer science, Kluwer, Dordrecht, 1997.

[9] O. Sjahputera, J. M. Keller, P. Mastakis, P. Gader, and J. Marjamaa, Histogram-based matching measures, Proceedings of the 19th International Conference of the North American Fuzzy Information Society NAFIPS'2000, Atlanta, Georgia, July 13–15, 2000, pp. 392–396.

[10] D. Weld and J. de Kleer, Qualitative reasoning about physical systems, Morgan Kaufmann, San Mateo, CA, 1989.