Cross-Stitching Using Little Thread[*]

Therese Biedl[†] Joseph D. Horton[‡] Alex Lopez-Ortiz[†]

Abstract

We consider the problem of cross-stitching a predetermined pattern on a piece of fabric. We show that computing stitching sequence that minimizes the amount of thread used when cross-stitching a pattern is NP-hard. However if the region to be stitched is connected, then the optimal solution can be obtained in linear time.

1 Introduction

Cross-stitching is an entertaining pastime where one "paints" pictures with needle and thread, usually on specially prepared cross-stitch canvases. See for example http://www.cross-stitching.com for many links for cross-stitching. Mathematically speaking, a cross-stitch pattern is simply a rectangular grid where some of the squares in the grid are filled with colors.

To stitch this pattern, we take another grid of holes (the dual of the first one, actually, though normally each grid-line consists of more than one thread of the canvas), and stitch a "cross" with a thread of the appropriate color for each square of the pattern. The colored thread is alternatedly above and below the grid of holes, and only the thread above the grid is normally visible. See Figure 1 for an illustration.

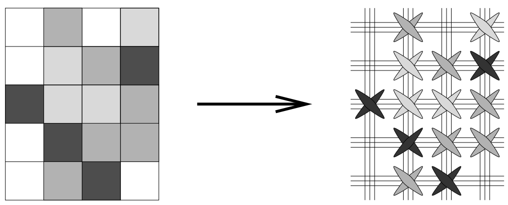

Figure 1: The pattern and the actual stitching.

In this paper, we consider the problem of stitching a pattern using as little thread as possible. To our knowledge, this problem has not been studied before. We consider the following "rules" for cross-stitching:

- If you end one stitch in one square, then you cannot start the next stitch in exactly the same square. For otherwise the thread would just "slip out" and the stitch would not hold.
- For each stitch, the lower-left-upper-right diagonal stitch, the *under-diagonal*, must come before the lower-right-upper-left-diagonal, the *over-diagonal*.[1] This is needed because switching the order of stitches results in a different reflection of light, thereby changing significantly the appearance of color in the resulting stitched picture.
- Arbitrarily long jumps behind the picture are not allowed. This is because during such jumps, the thread is either too tight (then the picture doesn't lie flat) or too loose (resulting in a tangled mess). If a long jump is required, then instead one should cut off the thread and make a knot and restart at the new place.

Of course, one would stitch the whole design by stitching one color at a time, and optimize the thread-use by optimizing it for each color. We hence study here only the problem of stitching a single color. Thus, we are given a *pattern* P, which is a set of stitches, and we want to find an order and direction of the diagonals of the stitches of P such that for each stitch the under-diagonal comes first in the order, the length connection between two consecutive diagonals is non-zero, and the total length of all connections (and tied knots) is minimal.

To compute length, we assume that the distance from one hole to a neighboring hole horizontally or vertically is 1, and the length of a stitch, being two diagonal distances, is $2\sqrt{2}$. Each diagonal of a stitch, other than the first one, is preceded by a stitch on the back side of length at least 1. Let k be the length required for a knot. Presumably this is longer than the distance between holes that are close together, so we assume than $k > \sqrt{2}$. Then the minimum length of thread required for n stitches is $2n(1+\sqrt{2}) + 2k - 1$. We say that a stitching of exactly this length is *perfect*. Most of our results are about perfect stitchings.

For many of our arguments, a parity argument will help. Let the holes of the grid be labelled by pairs of integers. The *parity* of a hole is the parity of the sum of

[*]Research supported by NSERC. The authors would like to thank the participants of an Algorithmic Problem Session at University of Waterloo, especially Steve Kalmar and Tomaš Vinář, for inspiring discussions.

[†]School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, {*biedl,alopezo*}*@uwaterloo.ca*.

[‡]Faculty of Computer Science, University of New Brunswick, Fredericton, NB E3B 5A3, *jdh@unb.ca*

[1]We don't really care which direction is considered "diagonal" and which one "off-diagonal", as long as it is the same for all stitches.

the coordinates. For each stitch, the parity of the holes used by either diagonal is the same, but the parity of the holes of different diagonals differ. Let the parity of a stitch be the same as the parity of the under-diagonal of the stitch. The stitches are partitioned into even and odd classes, so that neighbors are in different classes, like the black and white colors of a chessboard.

Each stitch is incident to four "holes" of the grid; we sometimes label these holes (when the stitch in question is clear) using the points of the compass as $h_{NW}, h_{SW}, h_{NE}, h_{SE}$. The endpoints of the under-diagonal of stitch s are h_{NE} and h_{SW}; we sometimes call these the *entry-points* of s since we must start stitching s at one of these two. The endpoints of the over-diagonal are h_{NW} and h_{SE}; we call these the *exit-points* of s.

2 NP-Hardness

The general problem of finding the shortest possible stitching of a pattern is NP-hard. We prove this by showing that a restricted version of the problem is equivalent to a known NP-complete problem.

Theorem 1 *The problem of determining whether a set of stitches can be stitched perfectly is NP-complete.*

Our proof of Theorem 1 starts with an apparent digression.

One variation of the problem is that we are required to do *one full cross at a time*, i.e., we must stitch the over-diagonal immediately after having stitched the under-diagonal. This still leaves the choice of the entry-point and the exit-point for each stitch, and in fact, for any choice of entry-point and exit-point this cane be done, leading to four possible ways of doing one full cross. See Figure 2 for an illustration. We mark the stitches with arrows to see from where to where the stitch is pulled.

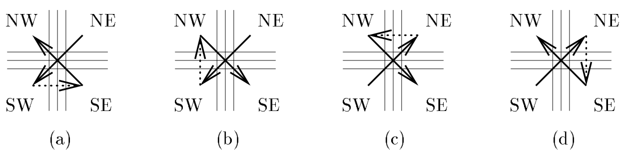

Figure 2: Four possible ways of completing one cross.

Consider the problem of finding a sequence that can be stitched perfectly, and in which the under-diagonal and over-diagonal are stitched immediately one after the other, for every stitch. When connecting stitches, the length of the connection must be 1, connecting the over-diagonal of the previous stitch to the under-diagonal of the following stitch. Hence the stitches must have the same parity. Also adjacent stitches in the sequence must be distance two apart in the Manhattan metric, so that adjacent stitches are either two steps apart horizontally or vertically, or diagonally connected.

We claim that any such sequence of stitches can be stitched perfectly, and one full cross at a time. Start the first stitch at an arbitrary entry point. For each stitch, choose the direction of the over-diagonal such that the exit-point that has unit distance to an entry-point of the next stitch in the sequence. This exists because the next point has distance two in Manhattan metric. Continue until all stitches are done.

On the other hand, for any perfect stitching of stitches of the same parity, the stitching must actually do one full cross at a time. Namely, in a perfect stitching we must alternate between an even diagonal and an odd diagonal. Since each under-diagonal must be done before its over-diagonal, there is only one over-diagonal that is ready to be done when an under-diagonal is completed, and that is the over-diagonal of the same stitch. We have proved:

Lemma 2 *A set of stitches that are all of the same parity can be run perfectly if and only if they admit a sequential ordering in which neighbors are two apart in the Manhattan metric. Moreover the stitches must be done one complete stitch at a time.*

Now we prove the theorem, by a reduction from Hamiltonian Path in a Grid Graph, which is NP-complete [3]. Let G be an arbitrary grid graph. For each vertex of G at coordinates (i, j), define a stitch with SW point at $(2i, 2j)$. Thus all stitches are even. From the lemma, these stitches can be run perfectly if and only if there is a they admit a sequential ordering in which the neighbors are two apart in the Manhattan metric. Such a sequential ordering corresponds exactly to a Hamiltonian path in G. This proves Theorem 1.

3 Stitching a contiguous region

In our NP-hardness proof, each stitch was isolated from all others. In contrast to this, we now prove that any connected region (by which we mean connected via edges; connected via a corner-point is *not* enough) has a perfect stitching pattern, and we can even impose some constraints on the location of the first and the last stitch.

Lemma 3 *Let P be a connected pattern and s be a stitch of P. For any entry-point p of s, and any exit-point q of s, there exists a perfect stitching pattern of P that starts at p and ends at q.*

Proof. The idea can be outlined as follows: perform a depth-first search of the graph of P starting at s. Stitch each under-diagonal while exploring a new edge of the depth-first search tree, and the corresponding over-diagonal while retreating along this edge.

To prove the claim more precisely, we use induction based on the number of stitches of P. If there is only one stitch s in P, then Figure 2 shows a suitable stitching. So now let there be more stitches in P. In the following we assume that $p = h_{SW}$ and $q = h_{SE}$; all other cases are symmetric (with a flip along one of the diagonals).

We stitch s as before, but "insert" a stitching of the rest of P inbetween. Since we can move from any even hole on s to any odd hole on s with a stitch of one unit length, it is not surprising that we can inductively stitch each part of the rest, beginning and ending with holes that also belong to s, and then move among the holes of s with unit length connections to come to the next area to be stitched. The precise description follows.

Note that $P - s$ may or may not be connected, but it breaks into at most four connected components. For $d \in \{E, N, W, S\}$, let P_d be the connected component of $P - s$ that is adjacent to s on the side of direction d. Note that some of these components may be empty or coincide with each other. We stitch the pattern as follows (see also Figure 3):

- Stitch the under-diagonal from $p = h_{SW}$ to h_{NE}.
- If P_N is non-empty: Stitch from h_{NE} to h_{NW}, and then stitch P_N starting at h_{NW} and ending at h_{NE}. (Note that h_{NE} and h_{NW} are an entry-point and exit-point of the stitch adjacent to s in P_N, so this is feasible by induction.)
- Now we are at h_{NE}. If P_E is non-empty and distinct from P_N: Stitch from h_{NE} to h_{SE}, and then p_E starting at h_{SE} and ending at h_{NE}.
- Now we are at h_{NE}. If P_W is non-empty and distinct from P_N and P_E: Stitch from h_{NE} to h_{NW}, then stitch P_W starting at h_{NW} and ending at h_{SW}.
- We're now at h_{NE} or at h_{SW}. If P_S is non-empty and distinct from P_N, P_E and P_W, stitch from the current location to h_{SE}, and then P_S starting at h_{SE} and ending at h_{SW}.
- We're now at h_{NE} or at h_{SW}. Do a unit stitch to h_{NW} and then the over-diagonal to $h_{SE} = q$. □

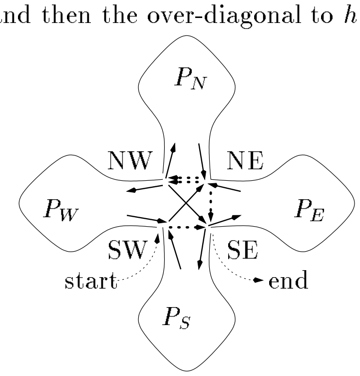

Figure 3: Stitching stitch s and its adjacent components, assuming all of them are non-empty and distinct. Dotted lines are stitches on the back.

We can obtain a perfect stitching of disconnected regions by using the above result on each component, as long as we can connect the components suitably. In particular, if we have a pattern of stitches that can be stitched perfectly, a full cross at a time, then we can insert stitching connected regions between them, using Lemma 3. By Lemma 2, we know when such a set exists.

Theorem 4 *Assume that a set of stitches can be partitioned into connected regions such that each region has a "root" stitch, and the set of root stitches admit a sequential ordering in which neighbors have distance 2 in Manhattan metric. Then the set of stitches can be run perfectly.*

The converse of this theorem is not true. The pattern in Figure 4 can be stitched perfectly, but the leftmost stitch has distance 3 in Manhattan metric to each of the other two, so there is no set of possible roots for the connected regions.

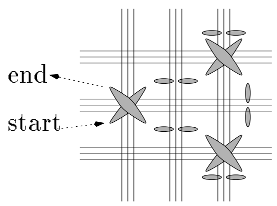

Figure 4: A counter-example for the converse of Theorem 4.

A second generalization of Lemma 3 concerns the beginning and the end. We can in fact start at an arbitrary stitch and end at an arbitrary (possibly different) stitch of the same parity, as follows. Assume that we have a connected set of stitches. Consider any two stitches of the same parity, say even parity, which we call s and t. Since the set of stitches is connected, these stitches are connected by a path of stitches that alternate in parity. This path can be found using depth first search in the graph. Let $n_1, \ldots, n_k$ be the even parity stitches on this path with $n_1 = s$ and $n_k = t$. By Lemma 2, these stitches can be run perfectly. The rest of the stitching pattern can be split into connected regions, each containing one of n_i, and so by Theorem 4 we can stitch the whole pattern perfectly.

Theorem 5 *Any connected set of stitches can be run perfectly, starting at any stitch, and ending at any stitch of the same parity. Moreover we can find this stitching in linear time.*

The requirement in Theorem 5 that the first and last stitch are of the same parity can not be removed: this must hold for any perfect stitching.

Theorem 6 *For any perfect stitching pattern, the parity of the first and last stitch is the same.*

Proof. Say the first stitch is even, so its under-diagonal is even, so the first diagonal that we stitch is even. The connection between two consecutively stitched diagonals has length 1, which implies that stitched diagonals alternate in parity. Since there is an even number of diagonals to be stitched, this means that the last stitched diagonal must be odd. But the last stitched diagonal must be an over-diagonal, so since it is odd, the stitch containing it is even. □

It is not hard to see that the stitching patterns for any of the constructive results in this section can be found in linear time.

4 Conclusion

In this paper, we show that the problem of finding the minimum amount of thread to cross-stitch a pattern is NP-hard, but it is easy to solve for connected patterns since there exists a perfect stitching.

The most interesting open problem is to find heuristics for stitching patterns that are not connected. If there are just a small number of connected components, one could solve the problem completely by looking at all the different ways to connect between the components. However we have no reasonable approximation algorithm, with performance ratio less than 2, say, if there are a large number of connected components. One should be able to do much better than this, particularly because the problem is in some sense geometric. One would think that one could find a reduction to, say, the Euclidean Travelling Salesman problem, but the apparently natural reductions are not symmetric. Can we instead use the underlying ideas of heuristics for the Euclidean Travelling Salesman problem (such as the Christofides heuristic [2] or the ε-approximation schemes [1, 4] and develop good approximation algorithms for the cross-stitching problem directly?

We do have one simple heuristic with a guaranteed performance bound, but which would cause laughter from a person who does cross-stitching. We could do each stitch independently, and tie a knot after each stitch. This gives an asymptotic performance ratio of $(2(k+\sqrt{2})+1)/2(1+\sqrt{2})$. In practice slightly better, but still laughable, would be to do a perfect stitching pattern for each connected region, and then tie a knot before moving on to the next.

If we drop the requirement that the under-diagonal must be stitched before the over-diagonal, then finding a stitching pattern with the least thread is easily reduced to the travelling salesman problem. Unfortunately, the triangle-inequality does not hold, so the standard approximation algorithms for TSP does not apply. Stitching one full cross at a time can also be phrased as a TSP problem, but again the triangle inequality doesn't hold.

Various other interesting problems come from further rules and variants often applied to cross-stitching:

- In our perfect stitching results (Section 3), we assumed that any unit stitch behind the pattern is allowed to be done. This is justified when the unit stitch is on the boundary of a cross (which in particular is the case for Lemma 3), but not always true otherwise.

  Specifically, to create a nice cross-stitch picture, no dark thread should run behind a light stitch, and no thread at all should run behind places where there is no stitch. This is needed because darker threads behind shine through in the final picture, giving undesired color effects.

  Thus, Lemma 2, and hence Theorem 4, only hold if all required unit distance connections are actually allowed to be done, because another (darker) color covers this area. What can be said about perfect stitching of non-connected patterns where this is not the case?

- Many cross-stitch patterns are more complicated than "each cross in one color". Some allow half-stitches, i.e., stitches that only go to the centre of the cross. Others require two colors for a cross, which means that one diagonal is stitched in one color and the other one in another color (it is not necessarily said which diagonal should be which.)

  A number of interesting variants arise from these modifications. In particular, consider the case that we only want to do the under-diagonal for each stitch. When do we have a perfect stitching pattern? This does not even exist for all connected patterns (though it does exist for patterns that have a Hamiltonian path.) Is this NP-hard to test?

References

[1] S. Arora. Nearly linear time approximation schemes for Euclidean TSP and other geometric problems. J. ACM, vol. 45, pp. 1-30, 1998.

[2] N. Christofides. Worst-case analysis of a new heuristic for the travelling salesman problem. Tech. Report, GSIA, Carnegie-Mellon Univ., 1976.

[3] A. Itai, C. H. Papadimitriou and J. .L. Szwarcfiter. Hamilton paths in grid graphs. SIAM J. Comput. vol. 11 no. 4, pp 676-686, Nov. 1982.

[4] J.S.B. Mitchell. Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric TSP, k-MST and related problems. SIAM J. Comput., vol 28, pp. 1298-1309, 1999.