

## 文書レイアウトにおける自動図表配置手法

4N-1

獅々堀正幹<sup>†</sup> 津田和彦<sup>††</sup> 青江順一<sup>†</sup>  
 (†徳島大学, ††住友金属)

### 1. まえがき

近年、DTP等の普及により、画面上で最終割り付けの確認をしながら文書を作成することが可能になった<sup>(1)</sup>。しかし、図表を文書中に配置するためには、図表の参照箇所、大きさ等を考慮しなければならない。しかも、利用者自身の配置が、適切か否かの評価も難しい。従って、あらかじめ与えられた配置基準に準拠した位置に図表を自動配置する手法は、文書編集システムに対する一つの重要な課題である。

本稿では、参照箇所の位置、配置位置の安定性、配置位置の均等性、文章と図表の関連性等の配置基準を定義し、これらの基準を満たす図表配置位置決定アルゴリズムを提案する。

### 2. 図表配置の基準とデータ要素の定義

#### 2.1 基準の定義

より適切な位置に図表を配置するための基準を、情報処理学会と電子情報通信学会の論文を対象に調査し、以下のように決定する。

番号順に配置する基準を**順序基準**、ページの上部または下部に配置する基準を**安定基準**、参照箇所の後方で、最も近い配置可能な領域に配置する基準を**後方基準**、図表と関連のある文章との距離を近くするための基準を**関連基準**、全体的にバランス良く図表を配置するために、参照箇所の前方向に図表を置くことも考慮し、一定の距離内（前後1ページ内）に図表を配置する基準を**均等基準**。また、論文では、始章、終章、参考文献等（以下、特殊見出しと略する）は特別な論理構造<sup>(2)</sup>として扱われており、その領域内には意図的に図表が配置されないため、これを**制約基準**と定義する。

#### 2.2 データ要素の定義

文書は段組、図表は1フレーム<sup>(3)</sup>幅の片幅図表と2フレーム幅の両幅図表を対象とする。

また、 $p$ 番目の段落の段落行数を $P\_LINE(p)$ ； $d$ 番目の図表の図表行数を $D\_LINE(d)$ ； $f$ 番目のフレーム内の利用可能な空き行数を $F\_AVAIL(f)$ ； $n$ 番目のページ内の利用可能な空き行数を $P\_AVAIL(n)$ で表す。更に、処理中の段落 $p$ と図表 $d$ を配置するフレーム番号をそれぞれ $PARA\_F(p)$ 、 $DIA\_F(d)$ で表し、図表 $d$ の参照箇所が存在するフレームを $REF\_F(d)$ とする。尚、段落の属性が本文の場合は、後方処理または均等処理；特殊見出しの場合は関連処理と制約処理；見出しの場合は関連処理が行われる。

### 3. 図表配置決定処理

#### 3.1 後方処理

##### 3.1.1 片幅図表配置処理

片幅図表の処理は、 $REF\_F(d)$ と $DIA\_F(d-1)$ の距離によって、図1に示すように分けられる。まず、 $REF\_F(d)$ が $DIA\_F(d-1)$ 以前にある場合（図1(a)）、 $DIA\_F(d-1)$ 以前のフレームに $D\_LINE(d)$ 以上の空き領域があったとしても、順序基準を満たすため、 $DIA\_F(d-1)$ 以前のフレームには配置しない。即ち、図表 $d$ は $DIA\_F(d-1)$ または $DIA\_F(d-1)+1$ の上部に配置される。また、 $REF\_F(d)$ と $DIA\_F(d-1)$ が同じ場合（図1(b)）、後方基準を満たすため、 $REF\_F(d)$ と $D\_LINE(d)$ が同フレームならば下部に、異なれば上部に配置する。 $REF\_F(d)$ が $DIA\_F(d-1)$ 以後にある場合（図1(c)）、 $REF\_F(d)$ 以前のフレームには空き領域がないので、図表 $d$ は $REF\_F(d)$ または $REF\_F(d)+1$ のフレームに配置される。

##### 3.1.2 両幅図表配置処理

両幅図表処理では、ページ内に両幅図表分（ $D\_LINE(d)*2$ ）の未使用行数が残っていたとしても、両フレームにそれぞれ $D\_LINE(d)$ の未使用行数がなければ配置できない。従って、ページ内に既配置の図表や文章の再配置処理が必要となるが、2つの制約が生じる。まず、既配置の図表を次ページに移動すれば、配置対象の図表が既配置の図表より前に配置されることになり、順序基準に反する。また、文章を次ページに移動すれば、配置対象の図表は対応する参照箇所より前に配置され、後方基準に反する。即ち、後方への再配置はページを越えてはいけないという一つの制約が生じる。次に、本手法は基準を満たしながら逐次的に処理するので、既配置の図表や文章を前方に再配置すれば、今まで満たしていた基準が破棄される。即ち、前方への再配置はできないという制約も生じる。これらの制約から、ページの最終フレームに含まれる図表や文章は再配置できないことになり、段組み文書では、再配置可能な図表や文章はページの先頭フレームのものに限られる。以上より、両幅図表処理は図2のように分類できる。

まず、ページの最終フレームに図表や文章が含ま

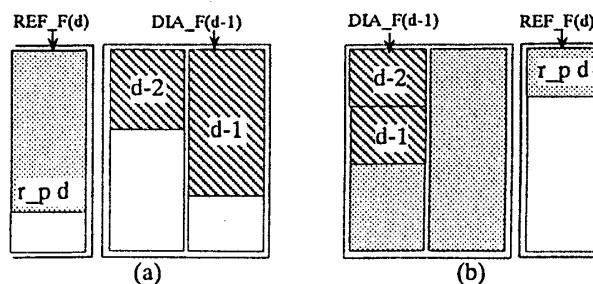


図1 片幅図表配置の説明図

A Method for Allocating Diagram on Automatic Document Layout

Masami Shishibori<sup>†</sup>, Kazuhiko Tsuda<sup>††</sup>, Jun-ichi Aoe<sup>†</sup>

<sup>†</sup>The University of Tokushima,

<sup>††</sup>Sumitomo Metal Industries, Ltd.

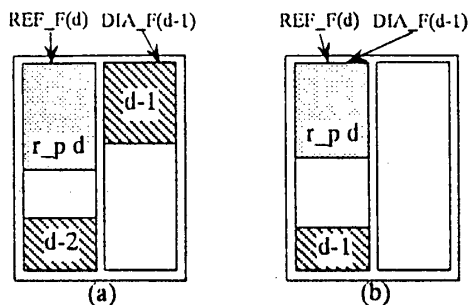


図2 両幅図表配置の説明図

れる場合(図2(a)), それらの図表や文章は再配置できないため, まず, 最終フレームの未使用行数を調べる。即ち, 最終フレームのF\_AVAILがD\_LINE( $d$ )より少なければ, 次ページの上部に配置する。また, D\_LINE( $d$ )の方が多ければ, 次にP\_AVAILを調べる。ここで, P\_AVAILがD\_LINE( $d$ ) $\times$ 2より多ければ, ページ内の図表や文章を再配置し, このページの下部に図表 $d$ を配置するが, 少なければ次ページの上部に配置する。次に, ページの最終フレームが空フレームの場合(図2(b)), 先頭フレームの未使用行数を調べる。即ち, 先頭フレームのF\_AVAILがD\_LINE( $d$ )より多ければ, そのページの下部に配置し, D\_LINE( $d$ )の方が少なければ, P\_AVAILを調べる。そして, P\_AVAILがD\_LINE( $d$ ) $\times$ 2より多ければ, ページ内の図表や文章を再配置し, このページの下部に図表 $d$ を配置し, 少なければ次ページの上部に配置する。

3.2 その他の処理

関連処理は, 話題の区切りに当たる見出し(区切り見出し)が出現した場合, 前章節内で参照された下配置済みの図表を上部に配置し直す。尚, 章節から章, 項から章節への変化に対しては話の流れが変化するので, 区切り見出しは, 章見出し, 及び項見出しの後に出現した節見出しとする。

多くの図表が集中した箇所では参照された場合, 図表が小さければ後方処理で対応できるが, 大きい場合は図表が文書の後方に追いやられて後方フレームに図表だけが固り, 図表と参照箇所との距離が大きくなる。均等処理では, 未処理の段落と図表の割合が閾値を越えた場合, 配置処理を後方処理から均等処理へ切り替える。また, 後方処理では参照箇所の後方に図表を配置していたが, 均等処理では参照箇所が出現したフレームに強制的に図表を配置する。

更に, 制約処理では, 制約領域以降に図表を配置する場合は, 制約領域が終了する位置を参照箇所とみなす。また, 制約処理以前の領域に図表を配置する場合は, 制約領域の行数を予め全段落行数から差し引くことにより, 均等処理が行われる。

4. 実験結果による評価

4.1 基準の評価

提案基準が実際の文書にどの程度適合するか論文50編(約440ページ, 図表556個, 片幅図表446個, 両幅図表110個)に対してを評価した結果を表1に示す。後方基準に不適応な図表は, 大半が均等及び制約基準を満たすために参照箇所の前

表1 評価Aによる結果

基準	適応率(%)	不適応な図表数(個)
順序	99.3	2
安定	100	0
後方	70.1	81
関連	91.4	24
均等	98.9	3
制約	100	0

表2 評価Bによる結果

基準	適応率(%)	不適応な図表数(個)
順序	100	0
安定	100	0
後方	100	0
関連	93.5	18
均等	98.9	3
制約	100	0

方に配置されたものである。即ち, 後方基準と均等基準は相反しており, 対象文書では参照箇所との距離を優先しているため, 後方基準の適応率が低下したと思われる。表1から全基準の平均的な適応率は93%となり, 導入した基準は妥当であると言える。

4.2 自動配置の実験結果による評価

4.1と同じ文書データに対して提案手法の配置位置を評価した結果を表2に示す。但し, 後方基準と均等基準は相反するため, 均等処理を行った図表は後方基準の対象外とした。均等処理が行われた図表は146個あり, もし, 均等処理をしなければ均等基準に反する図表は64個であった。この結果, 均等処理をしなければ適応率は88%であるが, この処理を行えば99%と向上し, 均等処理の有効性がわかる。表2から全基準の平均適応率は99%となり, 本手法の有効性が確認できる。

5. まとめ

本論文では, 適切な図表配置を行うための基準を定義し, この基準を満たす自動図表配置手法を提案した。また, 50編の論文データに対する実験結果から, 本手法による図表配置処理の有効性を実証した。片幅と両幅に属さない図表や他の段組みに対する処理の考案, また他種の文書についても本論文で提案した基準とその基準に基づく自動配置手法の有効性を確かめることが今後の課題である。

参考文献

(1)上林: “文書エディタの現状と将来展望”, 情処学会誌, 31-11, pp. 1535-1542 (1990-11).  
 (2)土井: “文書構造抽出技法の開発”, 信学論(D-II), J 76-D-II, 9, pp. 2042-2052 (1993-9).  
 (3)福井: “文書自動レイアウトシステムにおける図表配置候補生成方式”, 情処論文誌, 34-7, pp. 1499-1506 (1993-7).